

AD-A061 683

BOLT BERANEK AND NEWMAN INC CAMBRIDGE MASS
ON THE PREDICTION OF CONFUSION MATRICES FROM SIMILARITY JUDGMENTS--ETC(U)
OCT 78 D J GETTY, J A SWETS, J B SWETS

F/G 9/4

N00014-76-C-0893

UNCLASSIFIED

BBN-3719

NL

| OF |
AD
A061 683



END
DATE
FILMED
2-79
DDC

Bolt Beranek and Newman Inc.



12

AD A061 683

Report No. 3719

LEVEL

**On the Prediction of Confusion Matrices
from Similarity Judgments**

D.J. Getty, J.A. Swets, J.B. Swets, and D.M. Green



78 11 27 046

DDC FILE COPY

October 1978

Prepared for:
Engineering Psychology Programs, Office of Naval Research
ONR Contract No. N00014-76-C-0893
Work Unit No. NR 196-145

Approved for public release; distribution unlimited.
Reproduction in whole or in part is permitted for any
purpose of the United States Government.

Report No. 3719

Bolt Beranek and Newman Inc.

ERRATUM

The second reference on page 61 should be

Swets, J. A., Green, D. M., Getty, D. J., and Swets, J. B.
Identification and scaling of complex visual patterns. (Technical
Report 3536). Cambridge, Massachusetts: Bolt Beranek and Newman
Inc., November, 1977.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 14 BBN-3719	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER 9
4. TITLE (and Subtitle) ON THE PREDICTION OF CONFUSION MATRICES FROM SIMILARITY JUDGMENTS.		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
6. AUTHOR(s) David J./Getty, John A./Swets, Joel B. /Swets, and David M./Green		7. PERFORMING ORG. REPORT NUMBER 3719
8. PERFORMING ORGANIZATION NAME AND ADDRESS Bolt Beranek and Newman Inc. 50 Moulton Street Cambridge, Massachusetts 02138		9. CONTRACT OR GRANT NUMBER(s) 15 N00014-76-C-0893
10. CONTROLLING OFFICE NAME AND ADDRESS Engineering Psychology Programs Code 455, Office of Naval Research		11. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 196-145
12. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 12 T6p.		12. REPORT DATE October 1978
		13. NUMBER OF PAGES 74
		14. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) Approved for public release; distribution unlimited		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) signal detection, signal identification, human observers, complex visual stimuli, multidimensional scaling, decision processes in perception, pattern recognition		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) This study of human identification of complex visual stimuli is a step toward an integrated description of the perceptual process whereby stimuli are represented psychologically and the decision process that makes use of the stimulus representation in selecting an identification response. A multidimensional scaling procedure was applied to judgments of stimulus similar- ity to derive the dimensions of a perceptual space and the		

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

060 100

LB

Acknowledgment

We thank Jessie Kurzon and Barbara Freeman for their dedicated service as observers and Ms. Freeman for her able programming assistance.

Table of Contents

<u>Acknowledgment</u>	ii
<u>Introduction</u>	1
<u>The Multidimensional Perceptual Space</u>	2
<u>The Identification Model</u>	6
<u>Two Types of Identification Task</u>	12
<u>Method</u>	14
<u>Stimuli</u>	14
<u>Apparatus</u>	18
<u>Procedure</u>	19
<u>Similarity Task.</u>	19
<u>Complete Identification Task.</u>	19
<u>Partial Identification Task.</u>	20
<u>Observers</u>	21
<u>Derivation of the Perceptual Space: The Similarity-Judgment Task</u>	21

<u>The 8 x 8 Complete Identification Task</u>	23
<u>Results</u>	24
<u>Error Probability.</u>	24
<u>Confusion Matrix.</u>	24
<u>Model Analysis</u>	29
<u>Parameter Estimation.</u>	29
<u>Prediction of Confusion Matrices.</u>	29
<u>The 8 x 4 Partial Identification Task</u>	35
<u>Results</u>	36
<u>Error Probability.</u>	36
<u>Confusion Matrices.</u>	37
<u>Model Analysis</u>	45
<u>Parameter Estimation.</u>	45
<u>Predictions of Confusion Matrices.</u>	45
<u>Adaptive Tuning</u>	52
<u>Adaptive Tuning In Identification</u>	52
<u>Adaptive Tuning In Judgment of Similarity</u>	54

<u>Discussion</u>	55
<u>A Validation of MDS Procedures and the Decision Model</u>	55
<u>The MDS/Decision-Model Approach as an Aid in Predicting Identification Performance</u>	57
<u>References</u>	59
<u>Distribution</u>	62

Figure Captions

- Figure 1. Example of a stimulus configuration for which an asymmetric confusion matrix would be obtained (see text). 11
- Figure 2. Long-term spectra of eight underwater sounds (From Howard, 1976). 15
- Figure 3. Visual representations of the eight underwater sounds, photographed from the display monitor. 17
- Figure 4. 8 x 8 Experiment: Probability of a confusion error for each of the eight stimuli, for each observer. 25
- Figure 5. 8 x 8 Experiment: Distribution of response probability for each stimulus, for each observer. Obtained distributions are given by solid lines and filled circles; distributions predicted by the model are given by dashed lines and open circles. 27
- Figure 6. 8 x 4 Experiment, Condition 1: Distribution of response probability for each stimulus, for each observer. Obtained distributions are given by solid lines and filled circles; predicted distributions are given by dashed lines and open circles. 39
- Figure 7. 8 x 4 Experiment, Condition 2: Distribution of response probability for each stimulus, for each observer. Obtained distributions are given by solid lines and filled circles; predicted distributions are given by dashed lines and open circles. 41
- Figure 8. 8 x 4 Experiment, Condition 3: Distribution of response probability for each stimulus, for each observer. Obtained distributions are given by solid lines and filled circles; predicted distributions are given by dashed lines and open circles. 43
- Figure 9. Estimated salience weights on the four dimensions for each observer, for each condition in the 8 x 8 and 8 x 4 experiments (solid lines). The pattern of weights that maximizes probability of correct identification is given by the dashed line for each condition. 49

Table Captions

Table 1.	Psychological Coordinates for Each of the Eight Stimulus Patterns on Four Dimensions	22
Table 2.	8 x 8 Experiment: Matrix of Confusion Frequencies for Each Observer	26
Table 3.	8 x 8 Experiment: Estimated Parameter Values, and the Proportion of Variance Accounted for by the Model, for Each Observer	30
Table 4.	8 x 4 Experiment: Matrix of Confusion Frequencies for Each Observer in Each of the Three Conditions	38
Table 5.	8 x 4 Experiment: Estimated Parameter Values, and the Proportion of Variance Accounted for by the Model, Averaged Across Observers for Each of the Three Conditions	46

A reasonably complete account of the process by which humans are able to identify complex auditory or visual stimuli must address two issues: (1) the nature of the psychological representation of complex stimuli, and (2) the nature of the decision processes that act upon the internal representations to yield an identification response. While there has been substantial research directed at the representation and decision processes separately, there has been much less effort directed towards understanding the integration of these processes in complex-stimulus identification. Our major concern in this paper is with the relationship between the perceptual representation of complex stimuli and identification performance. The question is, given specific assumptions about the structure of the perceptual space, how well can we account for the pattern of responses observed in an identification task?

Our approach to the problem involves two parts: (1) the derivation of a multidimensional perceptual space for a set of complex stimuli from the application of a multidimensional scaling (MDS) procedure to judgments of stimulus similarity, and (2) the use of a probabilistic

decision model to predict the matrix of identification confusions from the geometric structure of the derived perceptual space. We are interested in the validation of the MDS procedure that would be supplied by a demonstration that the MDS-derived perceptual space can be used to predict behavior in an independent task. And we are interested in the possibility that the MDS procedure can be a substantial aid in understanding, and predicting behavior in, the fundamental task of stimulus identification.

We discuss first the rationale for, and assumptions made in, inferring a psychological space using MDS techniques. Then in the next section we describe our decision model for predicting identification confusions.

The Multidimensional Perceptual Space

Much of the work on the representation of complex auditory or visual stimuli suggests that perception is based on an analysis of the patterns along a number of psychological dimensions or features. In information-processing models, this processing is often referred to as "feature extraction" and is thought to reflect a selective reduction of information whereby perceptually important features are extracted from the pattern while other information is lost.

We may conceive of these dimensions as forming a multidimensional perceptual space in which each stimulus is represented as a point. This space, of course, is not directly observable. Both the set of dimensions comprising the space, and the loci of the stimuli within the space, must be inferred by indirect methods. Extensive development in recent years has led to the emergence of multidimensional scaling as an important method for deriving a representation of the perceptual space (e.g., Shepard, Nerlove, and Romney, 1972; Romney, Shepard, and Nerlove, 1972). MDS procedures are designed to decompose a matrix of pair-wise similarity judgments on a set of complex stimuli into a metric space of some (investigator-specified) number of orthogonal dimensions. Each stimulus is defined as a point in the space such that, ideally, the distances between pairs of stimuli in the space are monotonically related (inversely) to the degrees of judged similarity of the pairs.

The set of abstracted dimensions, and the relative loci of the stimuli within the space, may be interpreted to reflect the structure of the psychological space. This interpretation involves several assumptions. Most MDS procedures assume that the measure of the underlying psychological space is a member of the family of power

metrics.(1) This family includes the Euclidian and city-block metrics. The power metrics carry with them three important properties: (1) decomposability, (2) interdimensional additivity, and (3) intradimensional subtractivity (Tversky and Krantz, 1970). Decomposability means simply that the distance between any two points in the space is a function of dimension-wise contributions. Interdimensional additivity asserts that distance is a function of the sum of the dimension-wise contributions. Intradimensional subtractivity asserts that each dimension-wise contribution is the absolute value of the difference between the two points on that dimension. An intuitive implication of these properties is that, for a given pair of stimuli, the contribution to distance made by the values of the two stimuli on one dimension is independent of the values on all other dimensions.

Having obtained an abstract multidimensional solution, an investigator may attempt to relate the derived

(1) Power metrics are the class of metrics such that the distance between the points $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ is given by

$$d(x, y) = \left[\sum_{i=1}^n |x_i - y_i|^r \right]^{1/r}$$

for $r \geq 1$.

psychological dimensions to the known physical structure of the stimuli. Success in identifying the psychophysical functions relating psychological to physical dimensions is typically measured by a high correlation between values on a psychological dimension and values on the candidate physical measure, across stimuli.

MDS procedures have been used successfully to identify psychological dimensions underlying the perception of speech sounds (e.g., Pols, Van der Kamp, and Plomp, 1969; Klein, Plomp, and Pols, 1970; Shepard, 1972; Shaw, 1975); complex, non-speech sounds (e.g., Plomp and Stenneken, 1969; Miller and Carterette, 1975; Morgan, Woodhead, and Webster, 1976; Howard and Silverman, 1976; Howard, 1977); and complex visual patterns (e.g., Stenson, 1968; Shepard and Chipman, 1970; Hardzinski and Pachella, 1977; Pachella and Somers, 1978). In this context, success has usually meant that the derived multidimensional space accounts for a large proportion of the variability in the similarity judgments and that the revealed identity of the psychological dimensions is intuitively reasonable.

We suggest here that meeting either or both of these criteria does not provide strong evidence to support the validity of the derived representation. What is desirable is a demonstration that the perceptual space derived from

similarity judgments in one task can then be used to predict behavior in some other, independent task. Our present experiments provide an example of one such test, in that we use the MDS-derived space to predict performance in various identification tasks.

The Identification Model

We present here a decision model intended to predict the confusion matrix for a set of m stimuli in an identification task, on the basis of a multidimensional perceptual space. In the simplest case, the perceptual space is that revealed by application of an MDS procedure to judgments of similarity, and we shall confine our presentation in this section to that case. We note for later reference that the decision model can accept additional dimensions, as suggested by other evidence.

We take as our starting point the set of spatial coordinates, $\psi_{i,k}$ for each of m stimuli, s_i ($1 \leq i \leq m$), on each of n dimensions, d_k ($1 \leq k \leq n$), as provided by the MDS procedure. Though the model can be used with any MDS procedure that yields such a set of spatial coordinates, the data analyses presented in this paper are based on the INDSCAL procedure (Carroll and Chang, 1970; Carroll, 1972; Carroll and Wish, 1973). INDSCAL assumes that the judged

similarity between any pair of stimuli is a (decreasing) linear function of the Euclidean interstimulus distance in the underlying perceptual space. INDSCAL differs from most other MDS procedures in that it yields not only the spatial configuration of the set of m stimuli in the n -dimensional "group stimulus space," but also a vector of weights for each observer that reflects the relative importance or salience of each dimension for that observer. The effect of these salience weights for a given observer is to weight differentially the contribution of each dimension in determining interstimulus distance. Specifically, the distance between stimulus S_i and stimulus S_j , for an observer with salience weights w_k ($1 \leq k \leq n$) is given by

$$D_{i,j} = \left[\sum_k w_k (\psi_{i,k} - \psi_{j,k})^2 \right]^{1/2} \quad (1)$$

Our model will also assume the weighted Euclidean distance metric given in Equation (1). We will not assume, however, that the particular set of salience weights determined for each observer by INDSCAL in the similarity-judgment task necessarily applies to the identification tasks. In fact, we will show later that, within observers, the set of salience weights changes in predictable ways across different conditions of the identification task. Accordingly, the salience weights for each observer are

treated in the model as parameters whose values are to be estimated from the confusion data. We assume further that the salience weights are all positive ($w_k \geq 0$) and sum to 1 ($\sum_k w_k = 1$). The latter constraint is simply a normalizing convention which reflects the fact that only the relative magnitudes of the weights are meaningful in the model.

Having defined the set of interstimulus distances $D_{i,j}$, the next step is to relate these distances to interstimulus confusability in the identification task. On intuitive grounds, confusability should be some monotone decreasing function of interstimulus distance. We define a set of confusion weights $C_{i,j}$, assuming that confusability between stimulus S_i and S_j is given by

$$C_{i,j} = \exp(-aD_{i,j}), \quad (2)$$

where a is a sensitivity parameter, greater than 0. As a decreases towards 0, over-all stimulus confusability increases; as a becomes large, over-all confusability decreases.

Several consequences of this relation are worth noting: (1) with both a and $D_{i,j}$ bounded below by 0, $C_{i,j}$ is bounded between 0 and 1; (2) since $D_{i,j}$ is a distance $D_{i,j} = D_{j,i}$, and therefore $C_{i,j} = C_{j,i}$; and (3)

since $D_{ii} = 0$ for all i , $C_{ii} = 1$ for all i . The choice of this particular function, from several considered, was dictated by its clear superiority in accounting for our confusion data in preliminary analyses. It is the same assumption used successfully by Shepard (1957, 1958a, 1958b) in his work on stimulus and response generalization.

Finally, the conditional probability of giving the response assigned to stimulus S_j when stimulus S_i was presented is assumed to be the confusability of S_j with S_i relative to the summed confusability of all stimuli with S_i :

$$\Pr(R_j|S_i) = \frac{C_{i,j}}{\sum_k C_{i,k}} . \quad (3)$$

Equation (3) is essentially Luce's choice model (1963), with the added assumption that there are no differential response biases. While it would be a simple matter to include measures of response bias in the model, we have chosen to exclude them here for reasons of simplicity (fewer parameters to estimate) and because we have no reason to expect strong response biases. In our tasks, the a priori presentation probabilities (known to the observers) were equal across stimuli, the response set was homogeneous, and there were no differential payoffs.

We are also implicitly asserting in Equation (3) that the set of responses assigned to stimuli are sufficiently distinguishable that response confusions are negligible, and they are therefore not incorporated into the model.

It is relatively common in identification tasks to find asymmetries in the confusion matrix about the main diagonal; that is, it is often true that $\Pr(R_j|S_i) \neq \Pr(R_i|S_j)$. One well-known source of this asymmetry is response bias. Of interest here is the observation that there is a second possible source of confusion asymmetries, one which arises from the decision rule itself. This can most easily be seen in an example, shown in Fig. 1, in which three stimuli are embedded in a two-dimensional space. As drawn,

$$D_{1,3} > D_{1,2} > D_{2,3} > D_{i,i} = 0 \text{ (for all } i),$$

and therefore

$$C_{1,3} < C_{1,2} < C_{2,3} < C_{i,i} = 1 \text{ (for all } i).$$

Calculating $\Pr(R_1|S_2)$ and $\Pr(R_2|S_1)$, we find

$$\Pr(R_1|S_2) = \frac{C_{2,1}}{C_{2,1} + C_{2,2} + C_{2,3}} = \frac{C_{1,2}}{1 + C_{1,2} + C_{2,3}}$$

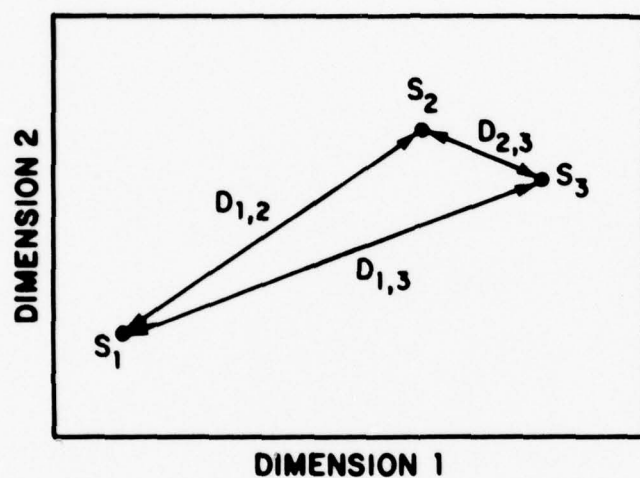


Figure 1. Example of a stimulus configuration for which an asymmetric confusion matrix would be obtained (see text).

and

$$\Pr(R_2|S_1) = \frac{C_{1,2}}{C_{1,1} + C_{1,2} + C_{1,3}} = \frac{C_{1,2}}{1 + C_{1,2} + C_{1,3}} .$$

Since $C_{1,3} < C_{2,3}$, we may conclude that $\Pr(R_2|S_1) > \Pr(R_1|S_2)$, thus demonstrating asymmetry in the confusion matrix about the main diagonal.

From this example, we can see that the confusion asymmetry between two stimuli S_i and S_j arises because the two response probabilities $\Pr(R_j|S_i)$ and $\Pr(R_i|S_j)$ are determined by the magnitude of the shared confusion weight, C_{ij} , relative to a sum of confusion weights, where the sum is determined from the "point of view" of the stimulus, S_i or S_j . These two sums may differ substantially in magnitude, depending upon the geometric configuration of the stimuli in the perceptual space.

Two Types of Identification Task

The primary test of the ability of the MDS-derived perceptual space, and the associated decision model, to predict identification behavior involves a straightforward identification task in which m responses are paired (one-to-one) with the m stimuli presented. We call this a "complete" identification task.

We also consider a "partial" identification task in which fewer than m responses are available for use with the m stimuli presented. This task corresponds to a situation in which some subset of the total number of stimuli present is of special interest -- the members of this subset constituting "signals" -- while the remaining stimuli are regarded as "noise," and as not requiring identification. We have referred to this class of task elsewhere as a "detection-and-identification" task (Swets, Green, Getty, and Swets, 1977). The observer is asked first to make a detection response ("signal" or "noise") and then to choose one of the available identification responses -- the one corresponding to the one of the signals most likely to be present.

The partial identification task employed here had three conditions, with a different subset of the m stimuli defined as signals in each condition. This variation across conditions reflects the practical fact that the subset of stimuli that is of special interest varies from one situation to another. The listener to degraded speech may want to concentrate at some time on distinguishing between just two particular phonemes, rather than among thirty or so. The sonar observer may wish to distinguish among ships and ignore variations in ocean depth, or vice versa. In the

present instance, the variation across conditions provides a test of our decision model's ability to deal with changes in the salience weights, which may reflect changes in the relative usefulness of the various perceptual dimensions as the set of signals is changed.

Method

Stimuli

Our stimuli consisted of visual representations of a set of eight underwater sounds, originally selected by Howard (1977) to represent a range of confusable natural and mechanically-produced sounds. They were referred to by Howard as (1) Sheet Cavitation (SC), (2) Biologics (BI), (3) Compressed Cavitation (CC), (4) Torpedo (TO), (5) Diesel Engine (DE), (6) Rain Squall (RS), (7) Steam Noise (SN), and (8) Flutter (FL). Their long-term energy spectra are shown in Fig. 2.

Our visual representations displayed the spectra as frequency (horizontal axis) versus time (vertical axis) versus energy (darkness -- the greater the energy, the darker the trace). We introduced periodicity as an additional physical dimension by sinusoidally varying the average darkness of the signal profile in the temporal

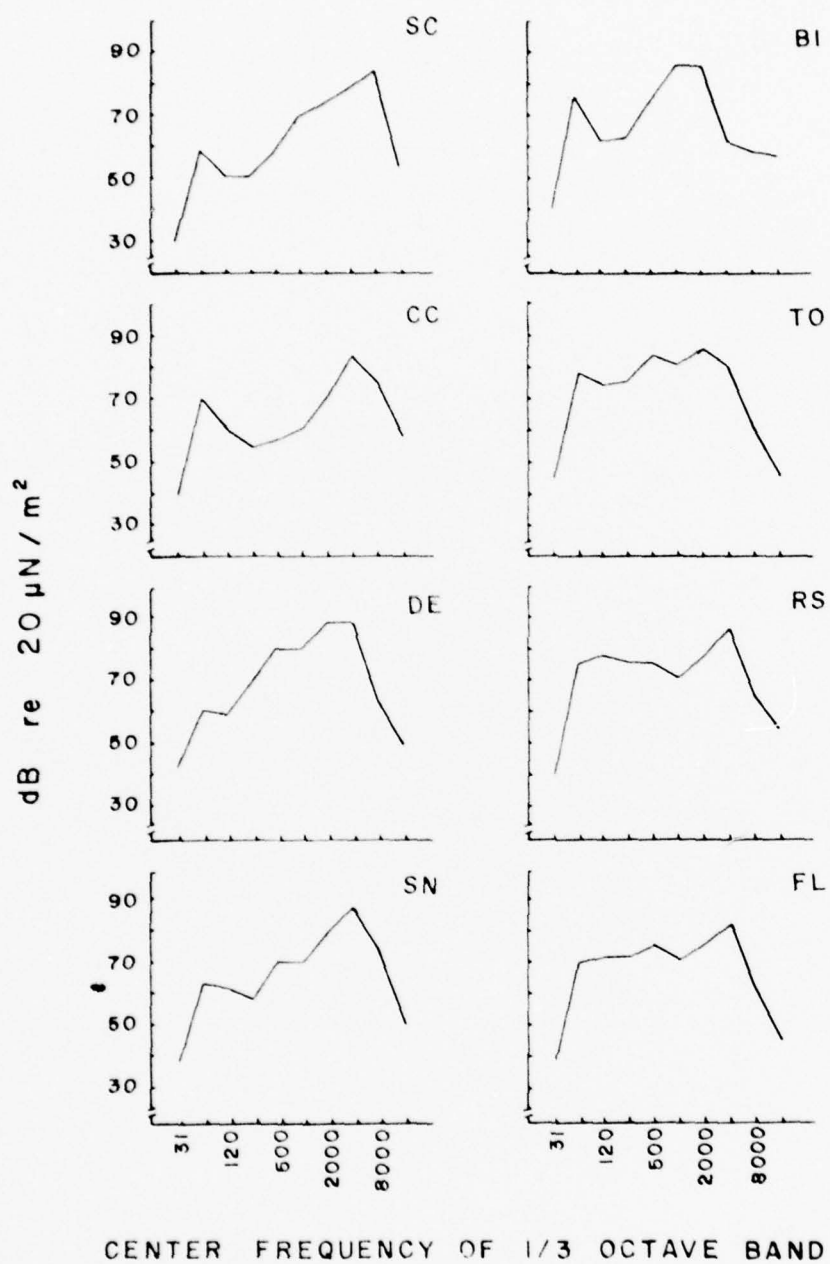


Figure 2. Long-term spectra of eight underwater sounds (From Howard, 1976).

(vertical) direction. We grouped (1) SC and (2) BI with (7) SN and (8) FL, giving all four relatively low-frequency periodicity, and gave the remaining four relatively high-frequency periodicity. Specifically, the cycles per stimulus were 15, 16, 17, and 18 for Stimuli 1, 2, 7, and 8, respectively, and 21, 22, 23, and 24 for stimuli 3, 4, 5, and 6, respectively. The resulting visual patterns are shown in Fig. 3.

The stimuli were constructed on a COMTAL model 8000-SA image-processing system, driven by a DEC PDP-11/34 minicomputer, and displayed in an area 24 cm wide by 12 cm high on a CONRAC 17-inch (43-cm) SNA television monitor. As is apparent in Fig. 3, we added a background of random noise to each stimulus pattern. The noise consisted of a 256 x 128 matrix of elements, each having an independent gray value sampled from a Gaussian distribution with mean 128 units and standard deviation 15 units on the 256 unit gray scale of the COMTAL.

Each stimulus pattern was constructed by subtracting from the noise background, sampled anew on each trial, the horizontal brightness profile corresponding to its long-term spectrum. Thus, increasing energy in the spectrum resulted in a darker trace. The spectral profiles of all eight signals were scaled to have the same space-average darkness

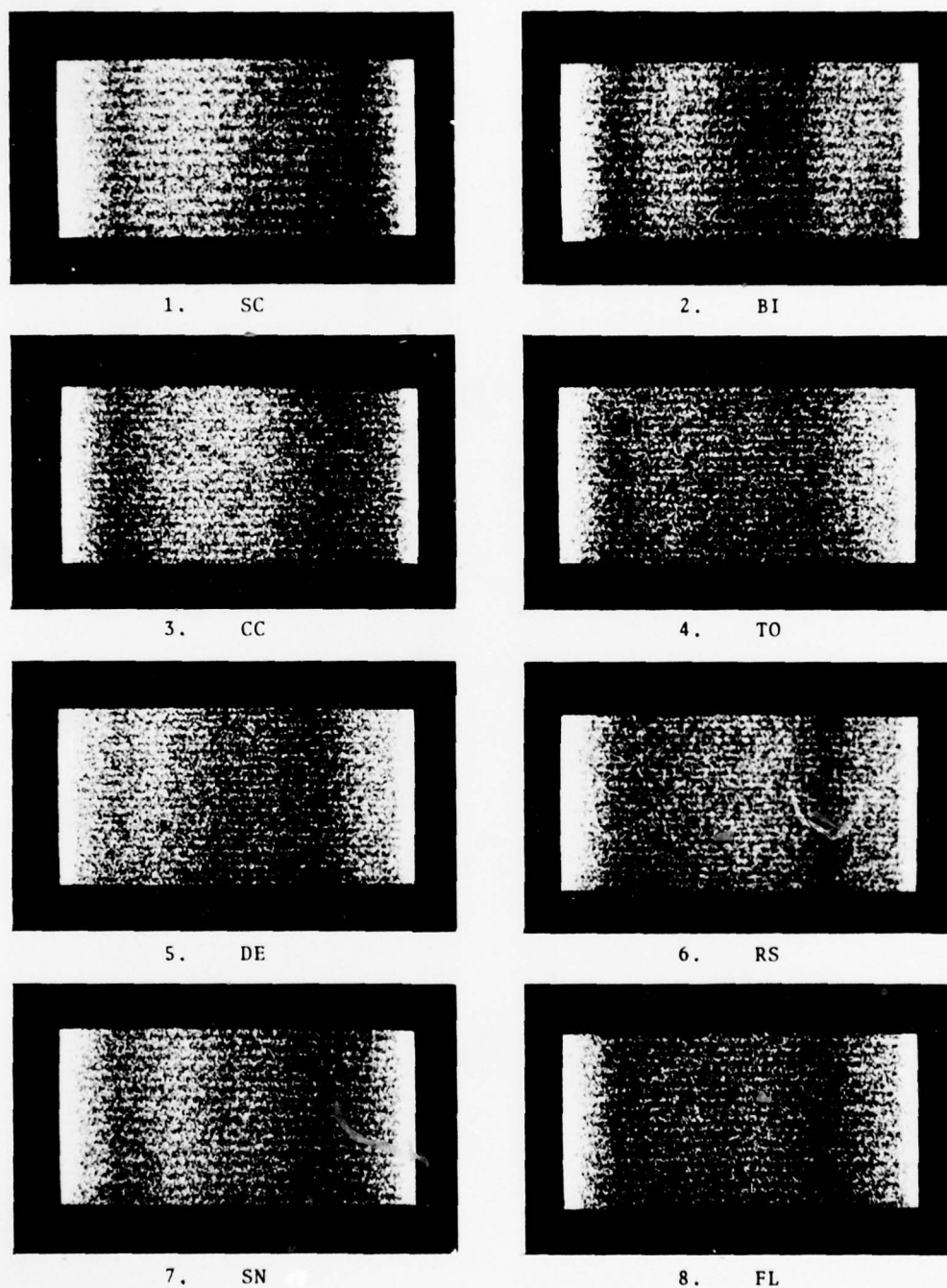


Figure 3. Visual representations of the eight underwater sounds, photographed from the display monitor.

of 20 gray units below the mean gray value of the noise background. In addition, the darkness of each point in a signal's profile was sinusoidally varied in the vertical dimension. For the similarity-judgment task, the peak-to-peak brightness variation was 200 percent about the steady-state value, resulting in high-contrast images similar to those shown in Fig. 3. For the various identification tasks, the peak-to-peak brightness variation was reduced to 60 percent about the steady state value, thereby reducing the image contrast considerably.

For all tasks, the brightness and contrast controls on the CONRAC monitor were adjusted such that the middle gray (128 units) had a luminance of about 62 cd/m^2 , and full white (255 units) a luminance of about 308 cd/m^2 .

Apparatus

In the various identification tasks, three observers sat at individual video computer terminals (Lear Siegler ADM-3A) approximately two meters from the stimulus-display screen, whose center was about 1.1 m above the floor. Ambient room lighting was maintained at a dim level.

All experimental events (stimulus display, response recording, and trial timing) were controlled by the PDP-11 computer.

Procedure

Similarity Task. The observers rated similarity of pairs of stimuli on a 10-point scale. Each of the 28 possible pairs was presented side by side by means of 35mm slides for 15 seconds, followed by a 15-second response interval. Each slide was a photograph of an high contrast display of one of the eight stimuli on the CONRAC monitor screen, similar to Fig. 3. Other details of the procedure are available elsewhere (Swets, Green, Getty, and Swets, 1977).

Complete Identification Task. Each trial began by blanking of the COMTAL screen, followed 2.5 seconds later by a 2.0 second low-contrast display of one of the eight visual stimuli. Each observer then made a self-paced identification response on the terminal keyboard, pressing one of eight keys labelled with the digits 1 to 8 (typing errors could be corrected with an "erase" key). Observers could make reference to a sheet on which were arranged labelled, high-contrast Polaroid photographs of the eight stimuli -- again, similar to those shown in Fig. 3. As each observer responded, the number of the presented signal was displayed on his/her terminal's screen. When all three had responded, the stimulus image was redisplayed along with the correct answer for about two seconds.

The sequence of stimuli was determined completely at random. Fifty trials were presented in a block, and three blocks were presented in a one-hour session.

Partial Identification Task. In this task, four of the eight stimuli were designated as "signal" and allowed as responses, while the remaining four were designated as "noise" and not allowed as responses. A different set of four stimuli was designated as the "signal" set in each of three conditions. The procedure for the partial identification task was generally the same as that described above for the complete identification task except that, for reasons discussed elsewhere (Swets, Green, Getty, and Swets, 1977), on each trial a stimulus pattern was presented in five sequential stages. At each stage, a successive fifth of the pattern was revealed, pushing down the display of earlier stages, and the observer made both a detection and an identification response. Our present analyses use only the identification responses from the last (fifth) stage of each trial when the complete display of the stimulus was in view. Observers were provided feedback -- either the signal number (when one of the four "signals" was presented) or the word "noise" (when one of the four "noise" stimuli was presented) -- at the end of each trial.

Observers

The three observers were members of BBN's technical staff, including one of the experimenters (JBS).

Derivation of the Perceptual Space:
The Similarity-Judgment Task

A set of 28 similarity judgments (all pairs of the eight stimuli) for each of the three observers was submitted to INDSCAL analysis. The psychological coordinates of the eight stimuli in the three-dimensional perceptual space revealed by INDSCAL are listed as $\psi_{i,1}$ to $\psi_{i,3}$ in Table 1. Details of the INDSCAL analysis are available elsewhere (Swets, Green, Getty, and Swets, 1977). Briefly, we found the three INDSCAL-derived dimensions to correlate highly with physical measures of "low-frequency energy," "mid-frequency energy," and "contrast," respectively.

Surprisingly, "periodicity" did not emerge as a psychological dimension, despite the fact that our observers reported it to be a salient dimension in the identification task. Moreover, as described in the earlier report, an INDSCAL analysis applied to the similarity judgments of another group of 14 judges yielded periodicity as a fourth dimension, along with the same three dimensions as given by the three observers.

Table 1

Psychological Coordinates for Each of the
Eight Stimulus Patterns on Four Dimensions

Stimulus	Psychological Dimensions			
	$\psi_{1,1}$ (low-frequency)	$\psi_{1,2}$ (mid-frequency)	$\psi_{1,3}$ (contrast)	$\psi_{1,4}$ (periodicity)
1	-.405	.427	.229	-.540
2	.441	-.143	.675	-.386
3	-.552	-.209	.243	.077
4	.380	-.083	-.344	.231
5	.317	.597	-.213	.386
6	-.154	-.521	-.233	.540
7	-.198	.207	.097	-.231
8	.171	-.276	-.454	-.077

In order to determine the gain, if any, in predicting confusion matrices when periodicity was included as a fourth dimension for the three observers, we created a periodicity coordinate for each stimulus. The assigned values, given in the last column of Table 1, are linearly related to our physical measure of periodicity, subject to the constraints -- used by INDSCAL in assigning coordinates on a psychological dimension -- that (1) the mean across stimuli is 0, and (2) the variance across stimuli is 1. We realize that the true psychological coordinates for periodicity are probably not linear with the physical measure; however, lacking a strong rationale for any other specific relationship, we may suppose that linearity is a good first approximation to the true relationship.

The 8 x 8 Complete Identification Task

There were eight response alternatives in the complete identification task, each corresponding to identification of one of the eight stimulus patterns. Of the 54 blocks of trials run, the first three were regarded as practice and omitted from analysis. In addition, the data of two other blocks, and part of a third, were lost due to equipment failures. The remaining 2421 trials for each observer were included in the analyses that follow.

Results

Error Probability. A plot of error probability against stimulus number, shown for individual observers in Fig. 4, reveals that the stimuli were not equally confusable. In fact, Stimulus No. 2 (BI) was never, or almost never, confused with any other stimulus. Individual observers showed similar patterns of errors across stimuli, as seen in Fig. 4. They also showed similar over-all error rates (13, 20, and 21 percent errors for observers BF, JK, and JS, respectively). The over-all probability of a confusion error, averaged across stimuli and observers, was 18 percent (1304 errors in 7263 trials).

Confusion Matrix. The matrices of raw confusion frequencies are given for each observer in Table 2. The patterns of confusions embedded in these numbers are most readily apparent when response probability distributions are plotted for each stimulus, shown separately for each observer by the solid lines and filled circles in Fig. 5. A prominent feature of these data is the high degree of similarity across the three observers in the confusion pattern for each stimulus. It is this matrix of confusion distributions that we seek to predict by the model, as described in the following section.

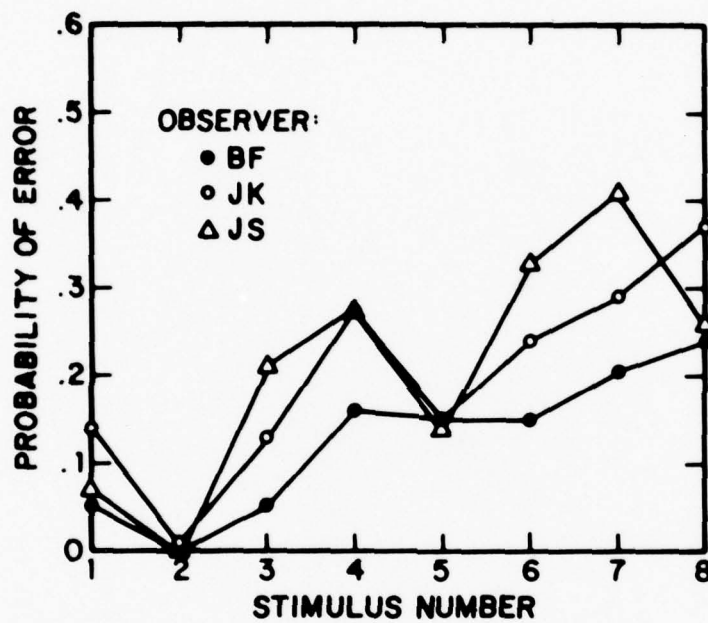


Figure 4. 8 x 8 Experiment: Probability of a confusion error for each of the eight stimuli, for each observer.

Table 2

8 x 8 Experiment: Matrix of Confusion
Frequencies for Each Observer

Observer	Stimulus	Response								Total
		1	2	3	4	5	6	7	8	
BF	1	273	0	1	1	0	0	12	0	287
	2	0	325	0	0	0	0	0	0	325
	3	2	0	271	0	0	3	8	2	286
	4	1	2	0	238	18	7	8	10	284
	5	0	0	3	26	249	8	3	3	292
	6	0	1	7	20	0	283	0	21	332
	7	23	1	5	1	3	1	245	29	308
	8	1	0	2	21	2	25	23	233	307
JK	1	248	0	1	0	1	0	36	1	287
	2	0	323	0	0	1	1	0	0	325
	3	0	0	250	2	0	26	3	5	286
	4	0	2	4	206	31	28	2	11	284
	5	0	0	0	17	247	24	0	4	292
	6	0	1	25	32	2	251	1	20	332
	7	12	5	9	3	7	4	218	50	308
	8	1	7	5	33	4	29	35	193	307
JS	1	267	0	1	0	0	0	17	2	287
	2	0	325	0	0	0	0	0	0	325
	3	0	0	226	11	1	36	9	3	286
	4	0	0	2	208	22	41	0	11	284
	5	0	0	3	27	250	7	2	3	292
	6	0	0	32	73	0	222	1	4	332
	7	23	2	8	9	6	1	181	78	308
	8	0	0	5	56	4	2	13	227	307

8x8 EXPERIMENT

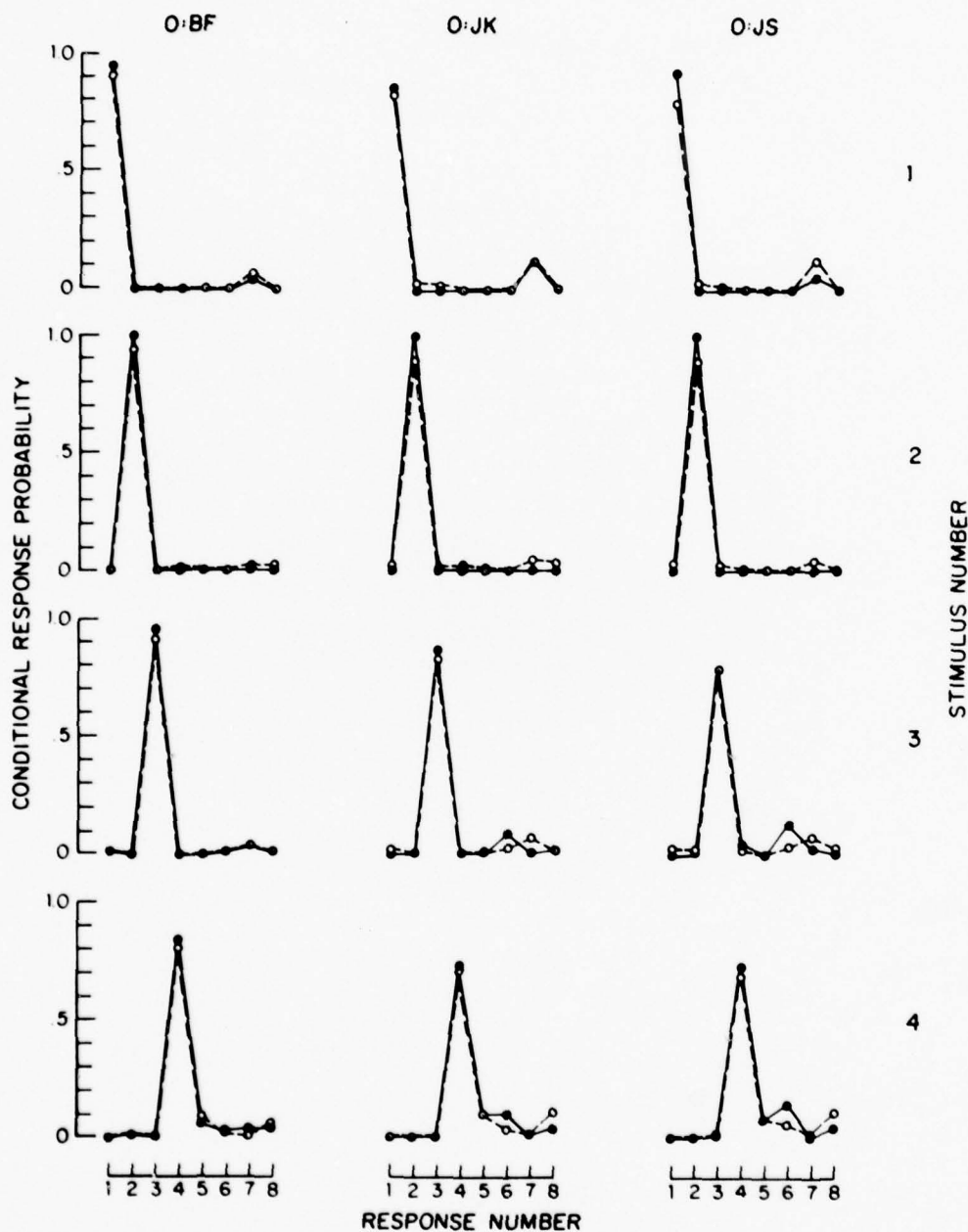


Figure 5A. 8 x 8 Experiment: Distribution of response probability for each stimulus (Stimuli 1 to 4), for each observer. Obtained distributions are given by solid lines and filled circles; distributions predicted by the model are given by dashed lines and open circles.

8x8 EXPERIMENT

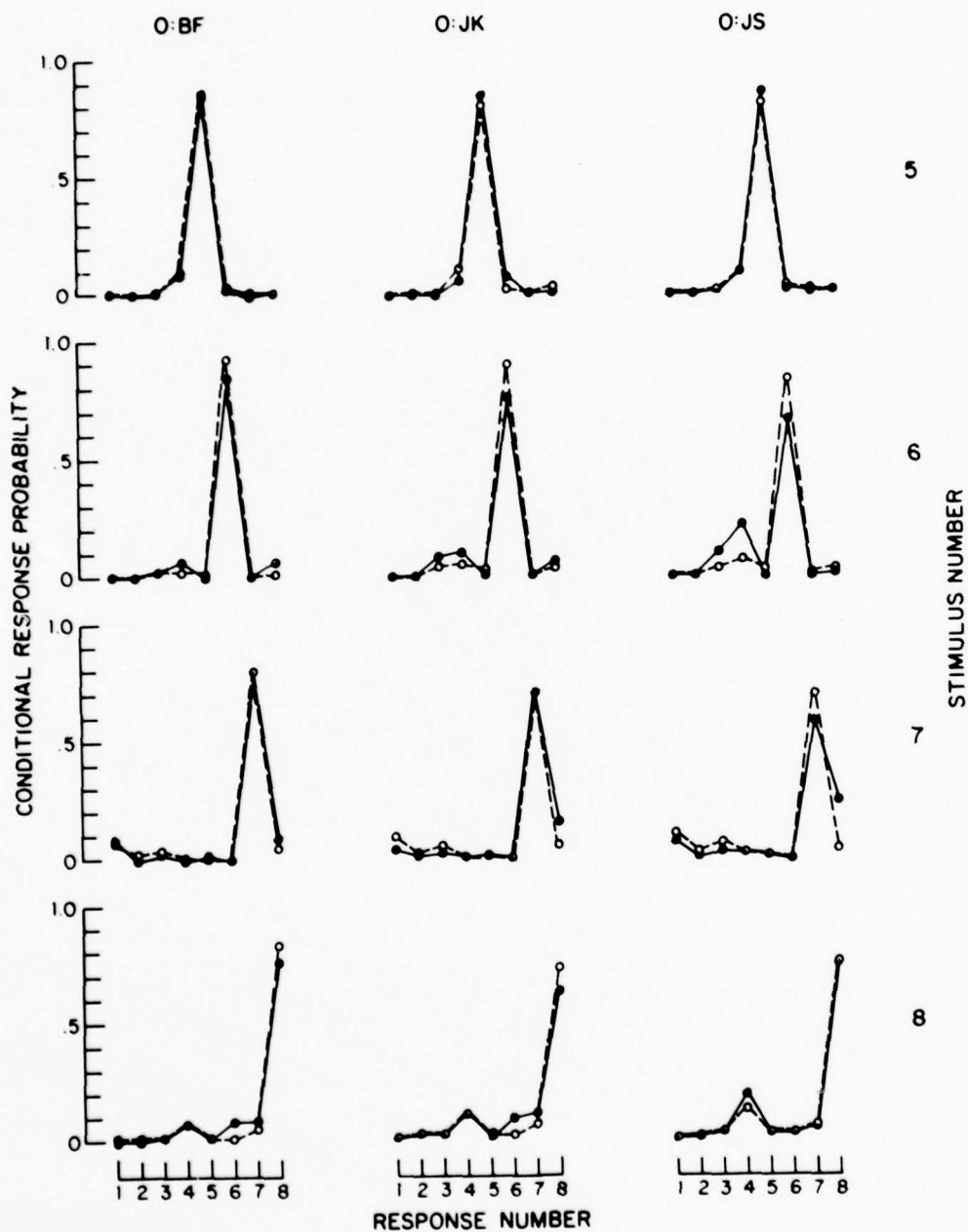


Figure 5B. 8 x 8 Experiment: Distribution of response probability for each stimulus (Stimuli 5 to 8), for each observer. Obtained distributions are given by solid lines and filled circles; distributions predicted by the model are given by dashed lines and open circles.

Model Analysis

Parameter Estimation. Model parameters were estimated separately for each observer using a negative-gradient hill-climbing technique that sought to minimize the sum of squared deviations between the predicted and obtained confusion matrices for that observer. Under several different analysis conditions regarding the dimension salience weights, discussed below, the number of parameters estimated ranged from one (a , the sensitivity parameter) to five (a and four salience weights, w_1 to w_4).

Prediction of Confusion Matrices. We first fitted the model to the individual confusion matrices of our three observers using the three dimension salience weights for each observer provided by the INDSCAL analysis. The parameter values of individual observers, and the proportion of variance accounted for in the data of individual observers, and the average values, are shown in Table 3A.

Estimating only a single parameter, a , for each observer, we found that we were able to account for 97 percent of the variance in the full confusion matrix, averaged across observers. On the other hand, if we included only identification errors (the off-diagonal elements) in the variance analysis, we accounted for only 21

Table 3

8 x 8 Experiment: Estimated Parameter Values
and the Proportion of Variance Accounted for by
the Model, for Each Observer

Obs.	Model Parameters					Proportion of Variance Accounted For	
	w_1	w_2	w_3	w_4	a	Full Matrix	Error Matrix

A. Three-dimensional model : w_1 - w_3 INDSCAL constrained

BF	.34*	.26*	.41*	--	10.9	.987	.281
JK	.57*	.29*	.14*	--	8.1	.971	.211
JS	.33*	.39*	.28*	--	7.7	.942	.126
AVG	.41*	.31*	.28*	--	8.9	.967	.206

B. Three-dimensional model : w_1 - w_3 free parameters

BF	.73	.14	.13	--	10.8	.988	.290
JK	.55	.27	.18	--	8.1	.972	.215
JS	.49	.28	.23	--	7.8	.943	.124
AVG	.59	.23	.18	--	8.9	.968	.210

C. Four-dimensional model : w_1 - w_3 INDSCAL constrained, w_4 free parameter

BF	.16*	.12*	.19*	.54	9.4	.991	.480
JK	.34*	.18*	.08*	.40	7.5	.978	.376
JS	.19*	.23*	.16*	.43	7.3	.955	.305
AVG	.23*	.18*	.14*	.46	8.1	.975	.387

D. Four-dimensional model : w_1 - w_4 free parameters

BF	.27	.08	.06	.59	9.5	.992	.543
JK	.30	.12	.11	.47	7.6	.978	.406
JS	.14	.14	.25	.47	7.4	.957	.325
AVG	.24	.11	.14	.51	8.2	.976	.425

*parameter not estimated; value derived from INDSCAL analysis

percent of the error-matrix variance. There are at least four reasons why the variance accounted for, considering only errors, may be small; we take them up in turn in the next four paragraphs.

First, the obtained relative frequencies calculated for the off-diagonal cells contain the effects of both quantization error (the observed frequency of each response is integral) and sampling error. Since the predicted probabilities of most off-diagonal cells are very small, the range of variation in the measures to be correlated is small. We may expect, on these grounds, that the unpredicted variation due to quantization and sampling error is substantial relative to the small total amount of variation present.

Second, the brief stimulus duration (2 seconds) may have resulted in some proportion of trials on which the stimulus either was not seen, or not seen long enough for adequate encoding. Responses on these trials should represent pure guesses, uniformly distributed across the set of responses (assuming no response biases). While guessing undoubtedly occurred occasionally, it is unlikely that it represents a major contribution to the confusion matrix. If it did, we would expect to see a non-zero baseline response frequency across all responses for a given stimulus, an

expectation which is not confirmed by the observed confusion matrices.

Third, the set of dimension salience weights derived by INDSCAL from the similarity-judgment task may not accurately represent dimension salience in the identification task. This possibility was examined by fitting the model with the salience weights, w_1 to w_3 , and the sensitivity parameter, a , free to vary. As seen in Table 3B, the estimated parameter values change somewhat from their INDSCAL-derived values; however, there is no significant increase in proportions of variance accounted for either in the full matrix or in the off-diagonal cells. Thus, for a model of three dimensions, the INDSCAL-derived salience weights are nearly optimal relative to the best possible performance of the model with unconstrained choice of parameter values.

This leads us to consider the fourth possible reason, namely, that one or more dimensions used by the observers in the identification task did not emerge in the INDSCAL analysis. As discussed earlier, we have good reason to suspect that the temporal periodicity present in the stimulus patterns is one such missing dimension. So we fitted the model to the confusion matrices a third time, using the INDSCAL-derived salience weights for each observer

for the first three dimensions, and allowing the salience weight for the fourth dimension (periodicity) to vary. By including periodicity as a fourth dimension, we found that the average proportion of variance accounted for in the full matrix increased from 97 percent to 98 percent, shown in Table 3C, and that accounted for in the off-diagonal cells increased from 21 percent to 39 percent, a substantial improvement. Moreover, with an average salience weight of .46 assigned to periodicity, it was by far the most important dimension among the four in determining interstimulus distance.

Finally, we wished to determine if any further improvement in prediction might be obtained by allowing the salience weights (constrained to their INDSCAL values in the last fit) to vary. The result, shown in Table 3D, was essentially no further increase in the proportion of variance accounted for. The values of w_1 , w_2 , and w_3 changed very little, on average, from the INDSCAL values -- indicating, as observed before, that the INDSCAL values were nearly optimal in terms of the model's ability to predict the full confusion matrix.

The pattern of results discussed above suggests that the effect of including periodicity as a fourth dimension is independent of, and additive to, the effect of freely

estimating the first three salience weights. Including periodicity as a fourth dimension increases the proportion of variance accounted for by 1 percent and 18 percent, respectively, in the full and error matrices when INDSCAL constraints are used (Tables 3A and 3C), and by 1 percent and 22 percent in the full and error matrices when parameters are freely estimated (Tables 3B and 3D). Freely estimating weights w_1 to w_3 increases the proportion of variance accounted for by 0 percent in both the full and error matrices when only three dimensions are used (Tables 3A and 3B), and by 0 and 4 percents when periodicity is included as a fourth dimension (Tables 3C and 3D).

We turn now from summary measures of goodness-of-fit to the prediction of individual cells in the confusion matrix, using the four-dimensional model with freely-estimated salience parameters. We choose this version of the model, in spite of the almost equally good fit noted above when INDSCAL constrained parameters are used, for consistency with data reported in the next section on the partial identification tasks, where fits of the different model versions were not equally good. The distributions of predicted conditional response probabilities are plotted for each observer in Fig. 5 as open circles connected by dashed lines, superimposed on the

obtained distributions. The difficulty one has in separating obtained and predicted curves attests to the considerable accuracy of the predictions.

There are occasional deviations between predicted and obtained probabilities that appear to be systematic, in that two of the three observers show the same pattern of deviations. For example when Stimulus 3 was presented, both observers JK and JS made Response 6 more frequently and Response 7 less frequently than predicted. This, and other such examples, may result from the observers' use of dimensions that were available in the set of patterns but not included in the model analyses. Over-all, though, the model that incorporates three INDSCAL-derived dimensions, and a fourth added, predicts quite well the individual cells of the confusion matrix.

The 8 x 4 Partial Identification Task

We may test the model in another way by considering how well it is able to predict the pattern of identification confusions when an observer is limited to responses associated with only a subset of the eight stimuli. In this section we apply the model to three conditions of an 8 x 4 partial identification task, in which only four of eight stimuli -- referred to as the "signals" -- correspond to

allowable identification responses, a different set of four stimuli in each of the three conditions.

The "signals" in Condition 1 were the four stimuli 1, 2, 5, and 6 of Fig. 3. These signals were not clearly distinguished from the remaining four "noise" stimuli on any of the four physical dimensions discussed previously. In Condition 2, the signals were stimuli 3, 4, 5, and 6, the patterns with relatively high-frequency periodicities. In Condition 3, the signals were stimuli 1, 3, 5, and 7, patterns which tended to have low values on our physical measure of mid-frequency energy.

The analyses that follow are based on 230 trials for each observer in each condition. On the average, each of the eight stimuli was presented about 28 times in each condition. An initial block of 30 practice trials has been omitted from analysis for each condition.

Results

Error Probability. Error rates can be defined only for the four "signals" in each condition. All responses on "noise" trials were necessarily errors since the noise stimuli did not correspond to allowable identification responses. Using only the signal subsets of the confusion

matrix for each condition, we found considerable variation in the error rates of the three observers, averaged across conditions: 5, 24, and 8 percent errors for observers BF, JK, and JS, respectively. There was also variation in the error rates of the three conditions, averaged across observers: 8, 21, and 9 percent for Conditions 1, 2, and 3, respectively. Finally, the over-all error rate in this experiment, 12 percent, was somewhat lower than that in the full 8 x 8 experiment (18 percent).

Confusion Matrices. The raw confusion matrices are given in Table 4 for each observer and for each condition. The corresponding response distributions are plotted in Figs. 6-8 for each stimulus, for each observer, and for each condition, by the filled circles connected by solid lines. As in the first experiment, the response distributions for a given stimulus and condition are generally very similar across the three observers. The most notable exceptions are the distributions for stimuli 6, 7, and 8 in Condition 2 (Fig. 7B), and for stimuli 2, 6, and 8 in Condition 3 (Figs. 8A and 8B). With the exception of stimulus 6 in Condition 2, these are all instances in which the identification response corresponding to the presented stimulus was not among the set of allowed responses, a fact we will return to shortly in considering the model's predictions.

Table 4

8 x 4 Experiment: Matrix of Confusion Frequencies
for Each Observer in Each of Three Conditions

Obs.	Stimulus	Condition														
		Response					Response					Response				
		1	2	5	6	Total	3	4	5	6	Total	1	3	5	7	Total
BF	1	28	0	0	0	28	32	0	0	0	32	32	0	0	1	33
	2	0	33	0	0	33	1	26	0	0	27	25	0	0	0	25
	3	9	0	1	18	28	25	0	0	1	26	0	30	0	2	32
	4	0	0	11	19	30	0	24	0	4	28	0	0	23	8	31
	5	0	0	27	1	28	0	2	21	0	23	0	0	23	0	23
	6	0	0	2	28	30	0	3	0	42	45	0	6	7	18	31
	7	23	1	2	1	27	16	1	1	2	20	2	1	0	27	30
	8	2	7	9	8	26	4	15	0	10	29	0	1	1	23	25
JK	1	28	0	0	0	28	28	0	3	1	32	29	0	0	4	33
	2	0	32	1	0	33	0	26	1	0	27	0	0	0	25	25
	3	4	0	3	21	28	20	0	0	6	26	1	26	1	4	32
	4	1	2	9	18	30	0	16	1	11	28	0	1	25	5	31
	5	2	0	20	6	28	0	2	14	7	23	0	1	22	0	23
	6	0	0	8	22	30	3	9	12	21	45	0	5	22	4	31
	7	18	1	3	5	27	5	1	8	6	20	4	1	1	24	30
	8	7	1	7	11	26	0	10	11	8	29	0	0	10	15	25
JS	1	28	0	0	0	28	32	0	0	0	32	29	0	0	4	33
	2	0	33	0	0	33	0	27	0	0	27	0	0	25	0	25
	3	1	0	0	27	28	24	0	0	2	26	0	32	0	0	32
	4	0	3	5	22	30	0	23	0	5	28	0	2	29	0	31
	5	0	0	22	6	28	0	1	22	0	23	0	0	23	0	23
	6	0	0	1	29	30	1	5	0	39	45	0	23	8	0	31
	7	14	5	4	4	27	14	5	1	0	20	1	3	0	26	30
	8	3	1	4	18	26	2	25	0	2	29	0	7	7	11	25

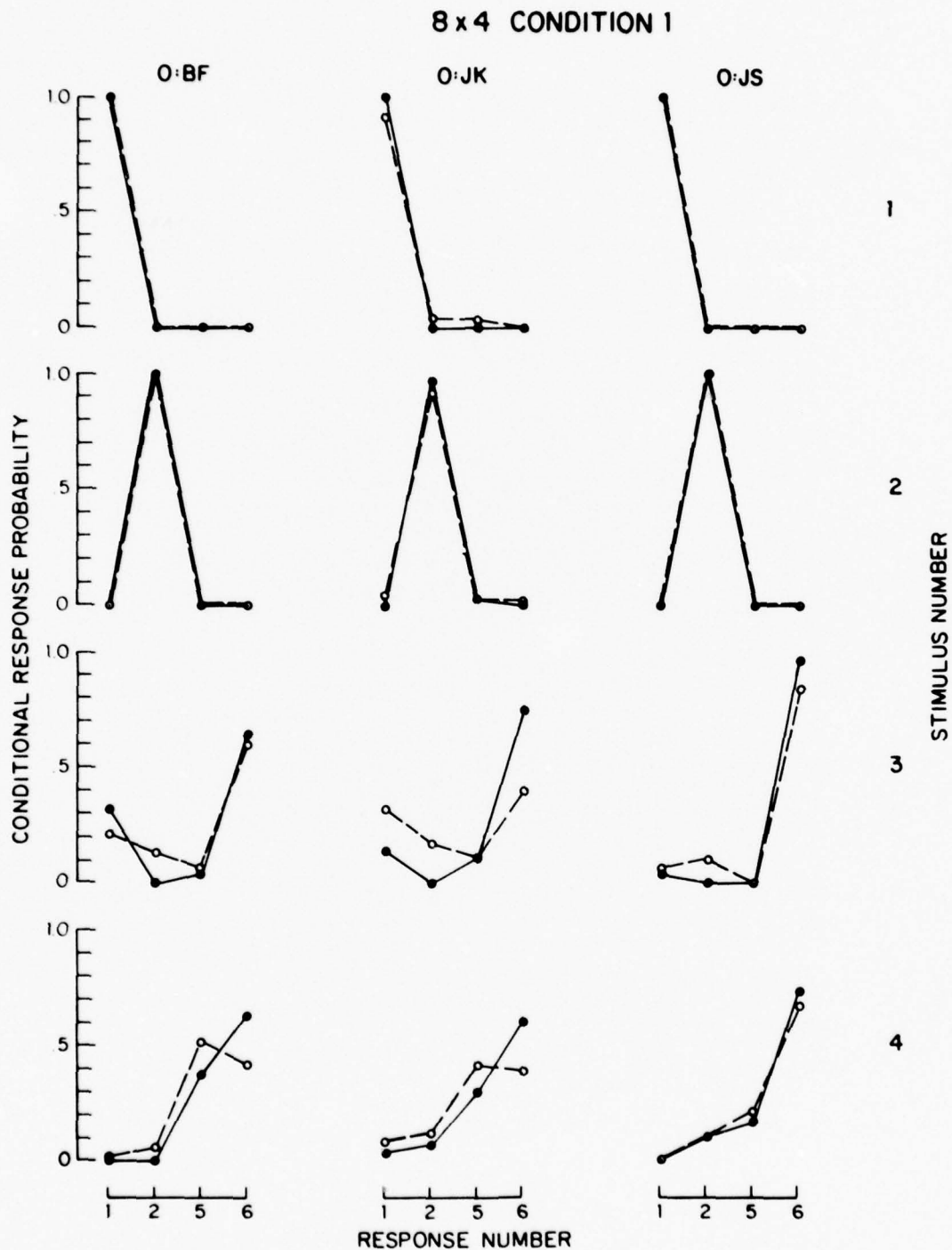


Figure 6A. 8 x 4 Experiment, Condition 1: Distribution of response probability for each stimulus (Stimuli 1 to 4), for each observer. Obtained distributions are given by solid lines and filled circles; predicted distributions are given by dashed lines and open circles.

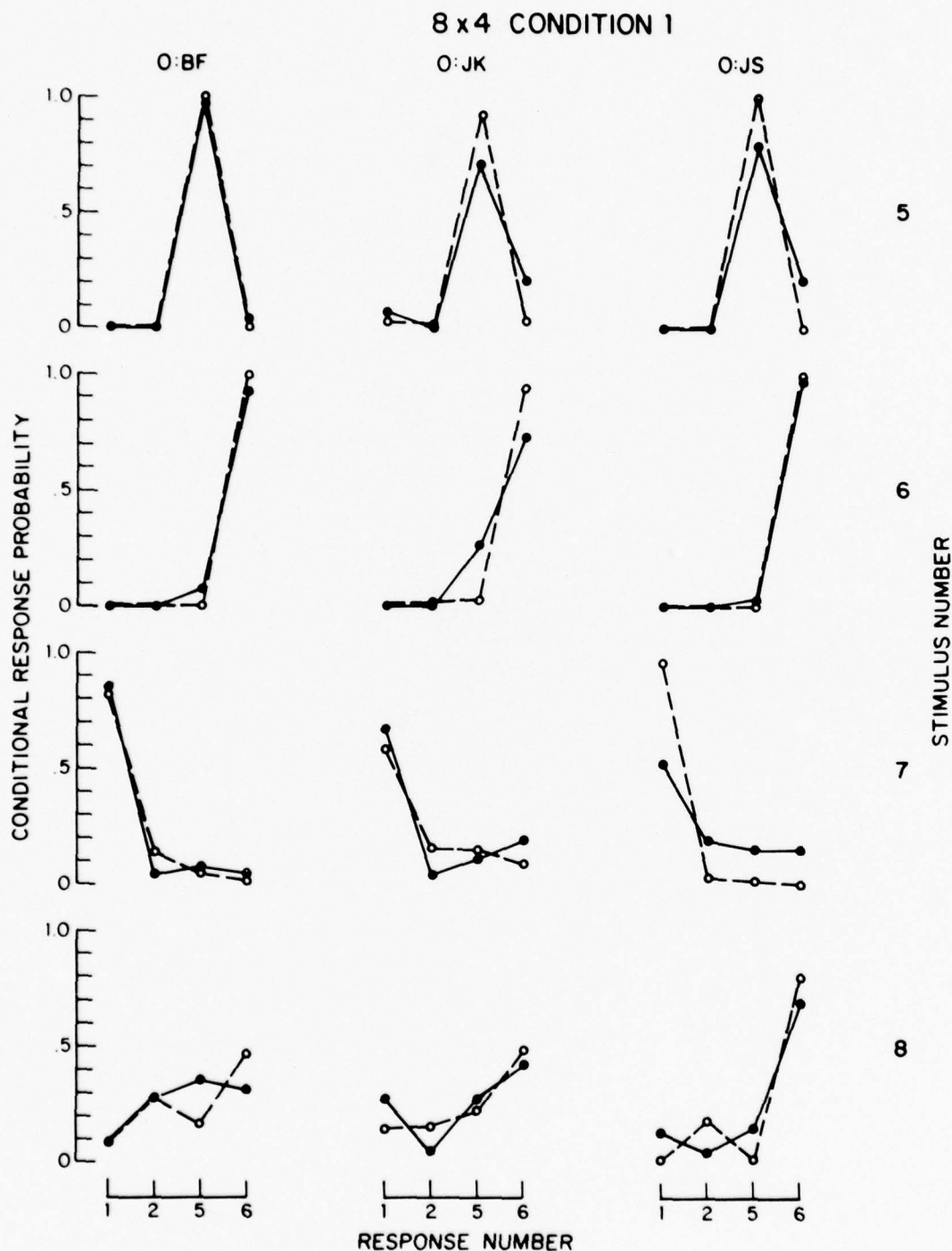


Figure 6B. 8 x 4 Experiment, Condition 1: Distribution of response probability for each stimulus (Stimuli 5 to 8), for each observer. Obtained distributions are given by solid lines and filled circles; predicted distributions are given by dashed lines and open circles.

Report No. 3719

Bolt Beranek and Newman Inc.

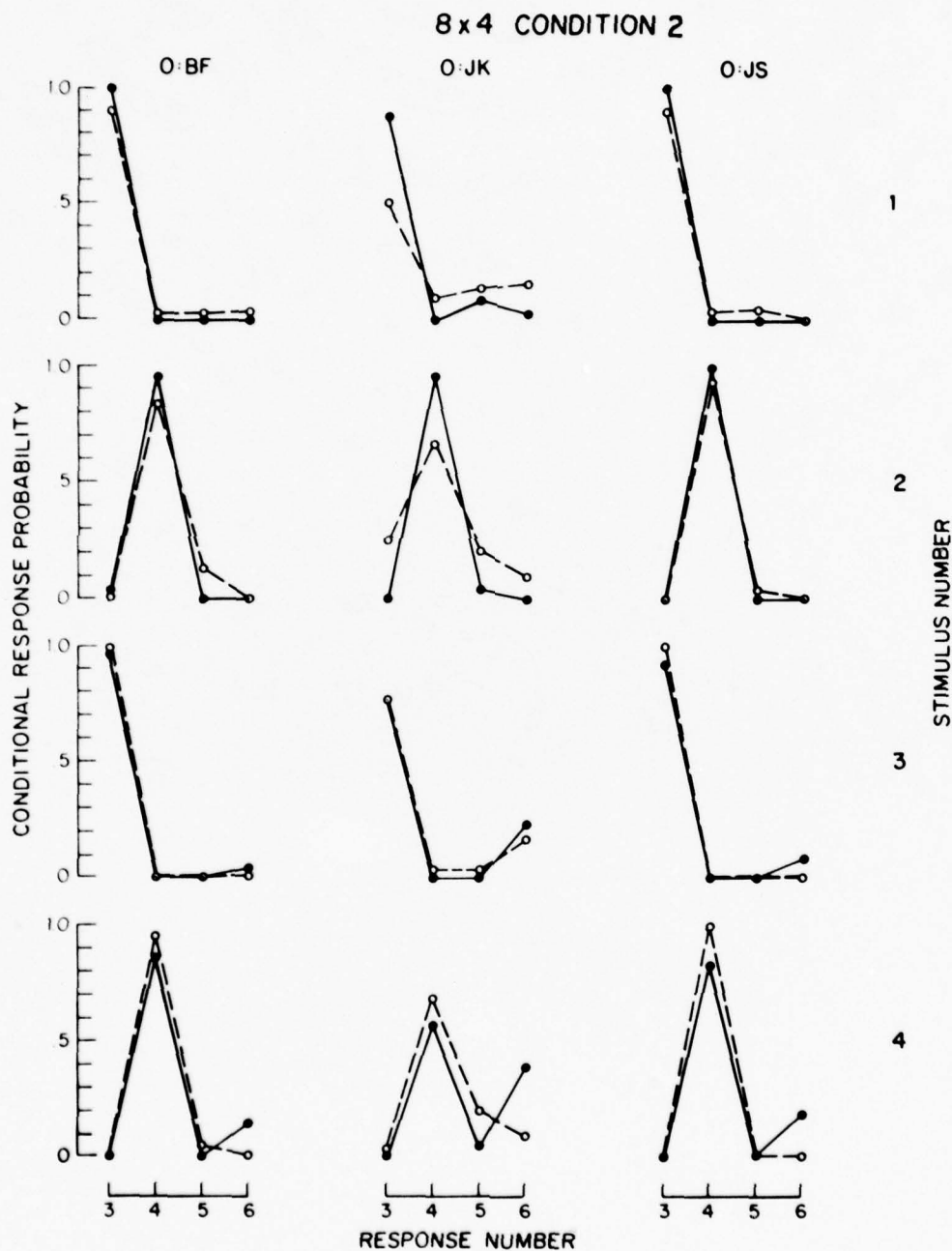


Figure 7A. 8 x 4 Experiment, Condition 2: Distribution of response probability for each stimulus (Stimuli 1 to 4), for each observer. Obtained distributions are given by solid lines and filled circles; predicted distributions are given by dashed lines and open circles.

Report No. 3719

Bolt Beranek and Newman Inc.

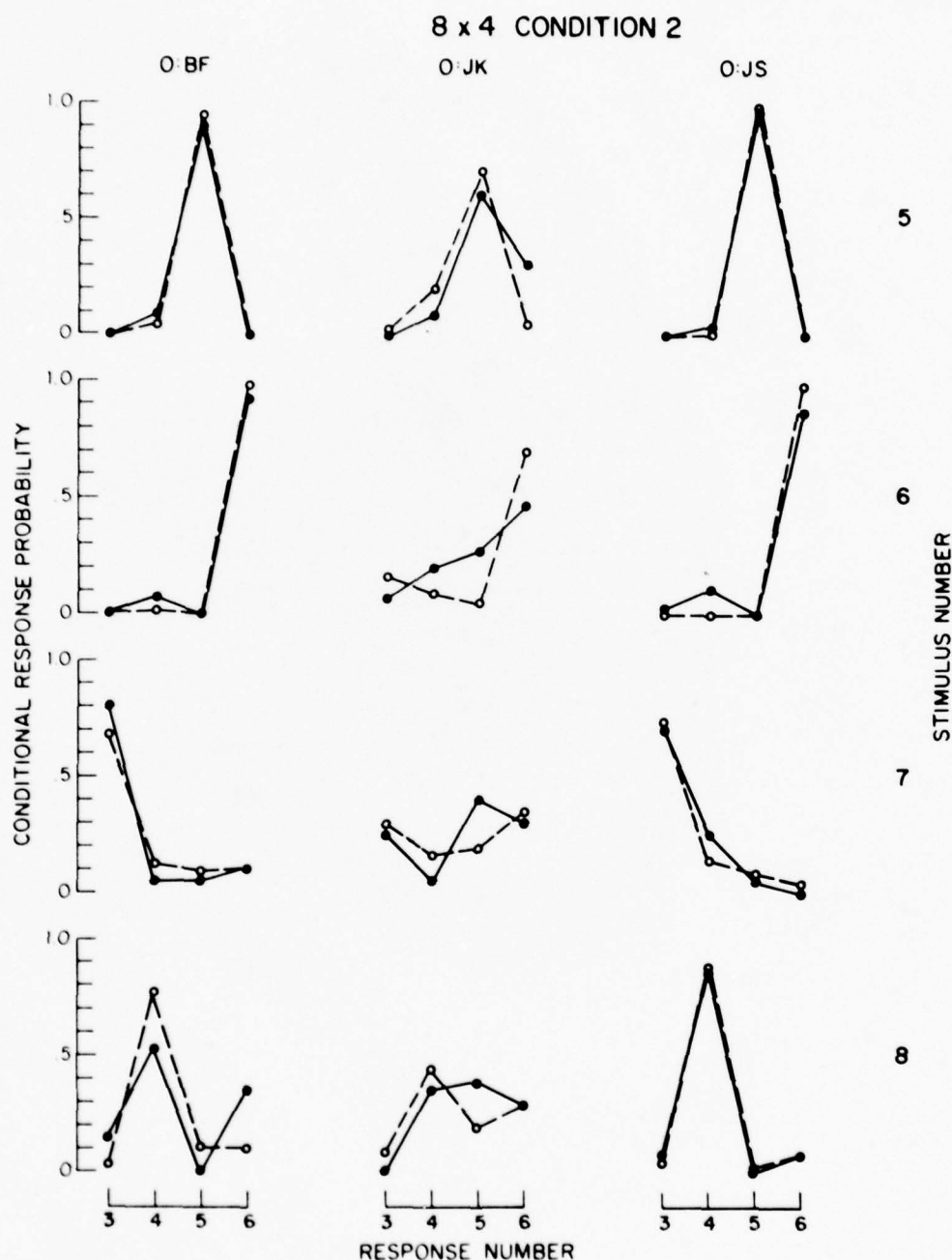


Figure 7B. 8 x 4 Experiment, Condition 2: Distribution of response probability for each stimulus (Stimuli 5 to 8), for each observer. Obtained distributions are given by solid lines and filled circles; predicted distributions are given by dashed lines and open circles.

8 x 4 CONDITION 3

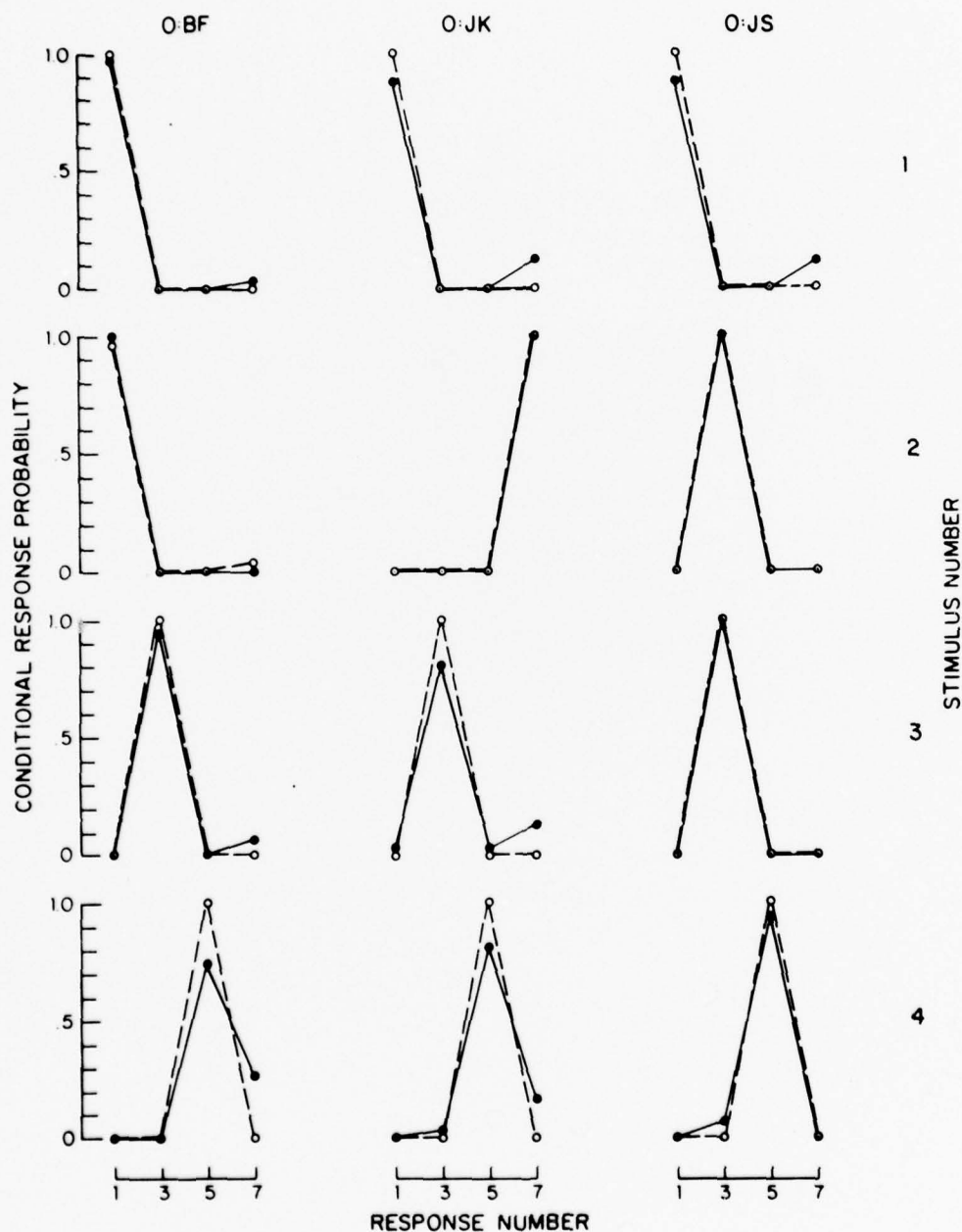


Figure 8A. 8 x 4 Experiment, Condition 3: Distribution of response probability for each stimulus (Stimuli 1 to 4), for each observer. Obtained distributions are given by solid lines and filled circles; predicted distributions are given by dashed lines and open circles.

Report No. 3719

Bolt Beranek and Newman Inc.

8x4 CONDITION 3

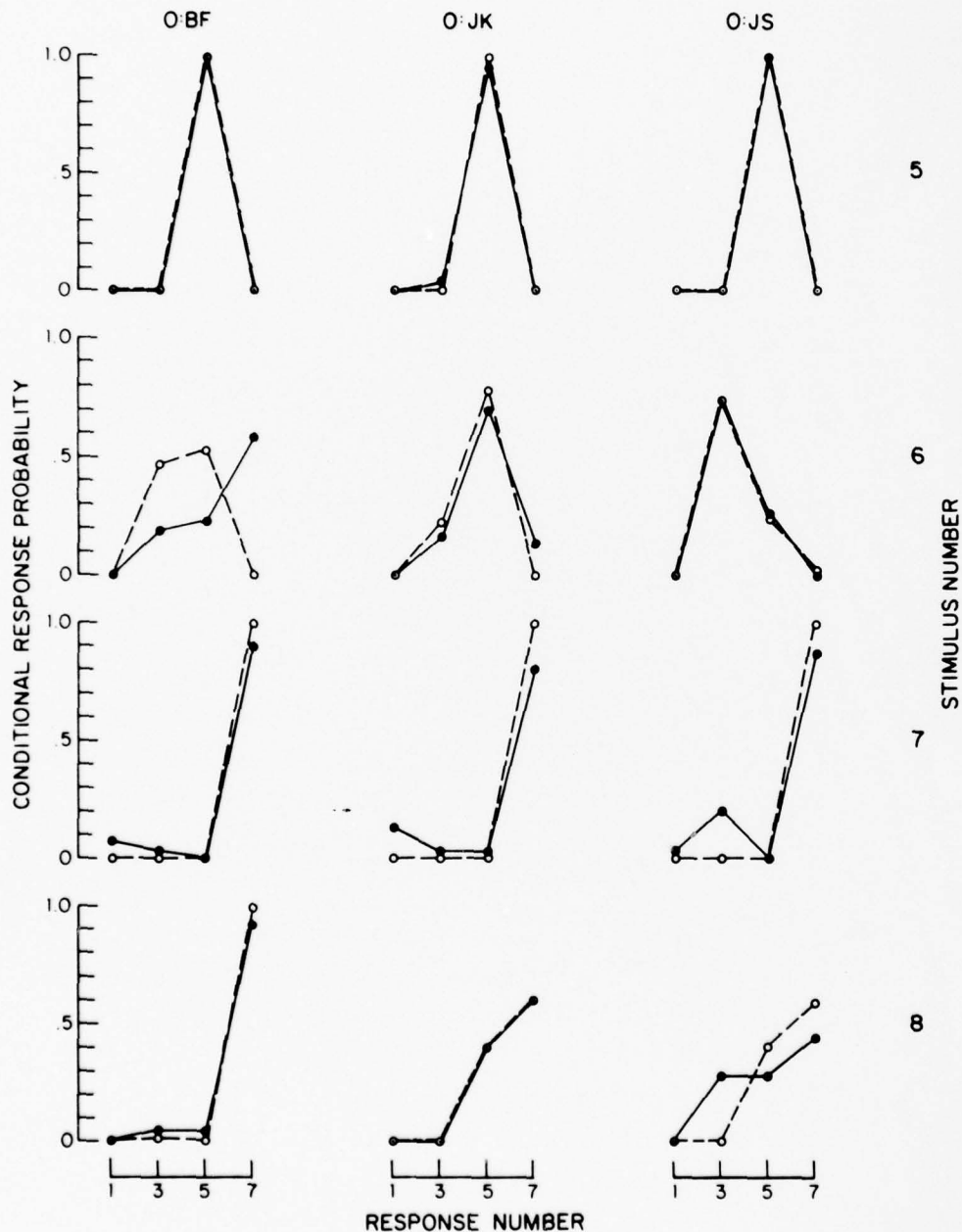


Figure 8B. 8 x 4 Experiment, Condition 3: Distribution of response probability for each stimulus (Stimuli 5 to 8), for each observer. Obtained distributions are given by solid lines and filled circles; predicted distributions are given by dashed lines and open circles.

Model Analysis

Parameter Estimation. The application of the model to the confusion matrices proceeded in much the same way as in the previous experiment. As before, the model was fitted to the data under four different conditions corresponding to combinations of two factors: (1) the inclusion or not of periodicity as a fourth dimension, and (2) the use of INDSCAL-derived or freely-estimated values for the three salience weights, w_1 to w_3 . In all cases, the psychological coordinates of the eight stimuli on the first three dimensions were taken from our INDSCAL analysis of the similarity-judgment data, and the coordinates of periodicity from our physical measure of that dimension. Model parameters were estimated separately for each observer in each of the three experimental conditions, and for each of the four versions of the model.

Predictions of Confusion Matrices. The parameter values and proportion-of-variance values for each condition, averaged across observers, and the average values across conditions, are shown for each of the four versions of the model in Table 5. Comparing the average proportion of variance accounted for by each model in Table 5 with the comparable average value for the 8 x 8 experiment in Table 3, it is clear that the model is accounting for somewhat

Table 5

8 x 4 Experiment: Estimated Parameter Values
and the Proportion of Variance Accounted for by
the Model, Averaged Across Observers for Each of the
Three Conditions

Cond.	Model Parameters					Proportion of Variance Accounted For	
	w_1	w_2	w_3	w_4	a	Full Matrix	Error Matrix

A. Three-dimensional model : w_1 - w_3 INDSCAL constrained

1	.41*	.31*	.28*	--	5.6	.814	.580
2	.41*	.31*	.28*	--	17.5	.772	.682
3	.41*	.31*	.28*	--	8.	.643	.322
AVG	.41*	.31*	.28*	--	10.7	.743	.528

B. Three-dimensional model : w_1 - w_3 free parameters

1	.31	.43	.26	--	10.6	.859	.693
2	.63	.33	.04	--	14.9	.859	.822
3	.44	.28	.28	--	16.7	.786	.672
AVG	.46	.35	.19	--	14.1	.835	.729

C. Four-dimensional model : w_1 - w_3 INDSCAL constrained, w_4 free parameter

1	.29*	.23*	.19*	.29	7.3	.855	.682
2	.39*	.30*	.27*	.04	17.5	.773	.684
3	.17*	.15*	.14*	.55	12.2	.746	.511
AVG	.28*	.23*	.20*	.29	12.3	.791	.626

D. Four-dimensional model : w_1 - w_4 free parameters

1	.20	.32	.17	.31	12.0	.897	.791
2	.58	.20	.00	.22	9.5	.858	.822
3	.32	.09	.21	.39	90.1	.937	.888
AVG	.37	.20	.13	.31	37.2	.897	.834

*parameter not estimated; value derived from INDSCAL analysis

less of the variance in the full matrix than before, but for substantially more of the variance in the error matrix. The first result is probably due in large part to the considerably smaller number of trials contributing to each of the confusion matrices -- approximately 30 -- resulting in less stable estimates of the observed conditional response probabilities. The second result follows from the inclusion of noise trials on which observers are forced to make confusion errors. As a consequence, the amount of variability available to be explained is large in the error matrices of the 8 x 4 experiment relative to the amount in the error matrices of the 8 x 8 experiment.

Paralleling the results of the analysis of the 8 x 8 experiment, the pattern of changes in the variance accounted for across the four versions of the model suggests that the effect of including periodicity as a fourth dimension is independent of, and additive to, the effect of freely estimating the first three salience weights. Including periodicity as a fourth dimension increases the proportion of variance accounted for by 5 and 10 percent in the full and error matrices, respectively (Tables 5A and 5C). Freely estimating the salience weights w_1 to w_3 increases the proportion of variance accounted for by 10 and 20 percent in the full and error matrices, respectively (Tables 5A and

5B). Doing both -- including periodicity as a fourth dimension and freely estimating the salience weights w_1 to w_3 -- increases the proportion of variance accounted for by 16 and 30 percent in the full and error matrices, respectively (Tables 5A and 5D).

The substantial improvement in the model's predictions, when the salience weights w_1 to w_3 are freed of their INDSCAL-derived values and independently estimated in each condition, is in contrast to the lack of improvement observed in the 8 x 8 experiment. The reason is apparent in Fig. 9, which shows the estimated salience weights (including the periodicity dimension) for each observer, separately for the 8 x 8 experiment and each of the three conditions of the 8 x 4 experiment. The pattern of estimated weights clearly changes from one condition to another, suggesting that the observers modified their set of salience weights from condition to condition according to the composition of the set of four stimuli defined as signals. In Conditions 1 and 2, the patterns of weights appear quite similar across observers; in Condition 3, they appear relatively different.

The predicted response distributions for each stimulus and for each observer in each of the conditions are shown by the open circles connected by dashed lines in Figs.

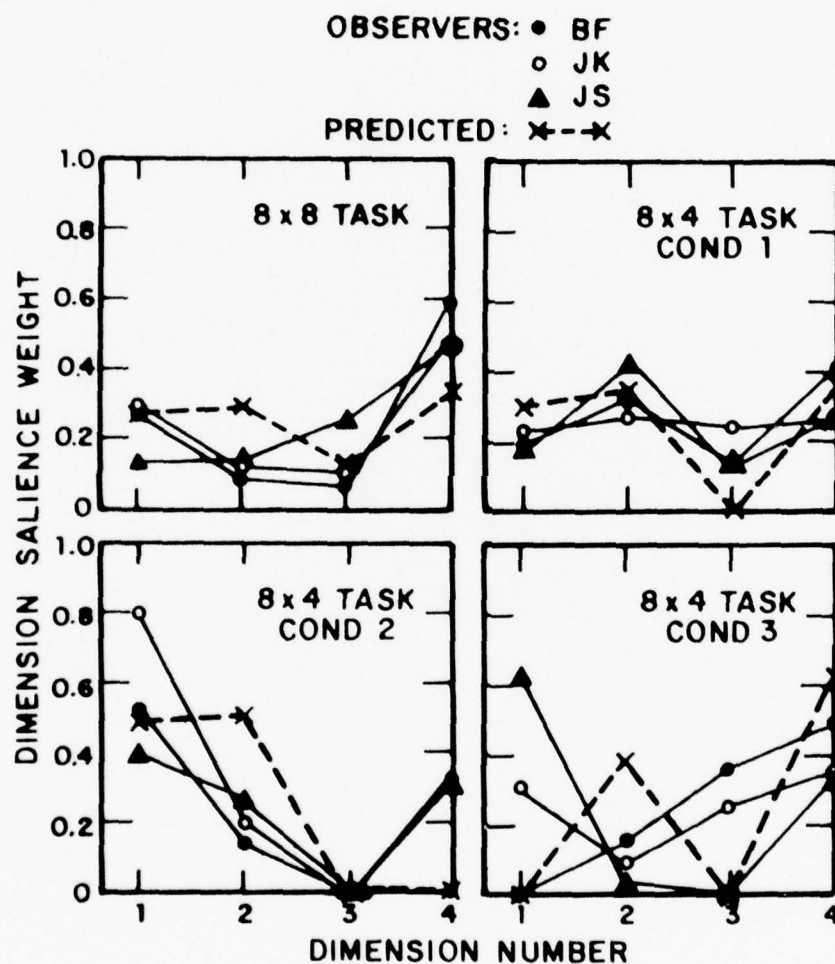


Figure 9. Estimated salience weights on the four dimensions for each observer, for each condition in the 8 x 8 and 8 x 4 experiments (solid lines). The pattern of weights that maximizes probability of correct identification is given by the dashed line for each condition.

6-8. These predictions are based on the four-dimensional model with freely estimated salience weights. Scanning over the large number of distributions, the over-all impression is that there is a remarkably good agreement between predicted and obtained distributions. It is worth noting that a ten-percent deviation between obtained and predicted probabilities corresponds to a difference of only about three responses, given that each obtained distribution is based on about 30 trials.

We mentioned earlier that most of the cases, in which the observed response distributions for a given stimulus differed among the three observers, were instances of noise stimuli. We now observe that most of the cases in which there is a substantial deviation between predicted and obtained response distributions are also instances of noise stimuli (e.g., Stimulus 6 for Observer BF in Condition 3). In fact, the structure of the model provides an insight into the source of these deviations. It will sometimes be true that the interstimulus distances between a particular noise stimulus and each of the four signal stimuli allowed as responses are all very large, and thus that the corresponding confusability weights are all very small. The observer would have no difficulty in rejecting all four responses as corresponding to the true "identity" of the

presented stimulus. In this predicament the observer most likely reinterprets the task to respond with the stimulus among the allowable four which is "most similar" to the presented stimulus, even though it is clearly incorrect. However, if the confusability weights for all four responses are extremely small, we may expect the observed response distribution to be very sensitive to small differences in salience weights, and also to other decision processes.

A particularly good illustration of the problem is found in the response distributions to noise Stimulus 2 in Condition 3 (Fig. 8). Each observer made a single response exclusively -- but a different response for each observer. In all likelihood, each observer was aware that the pattern was Stimulus 2 at each presentation. Having decided once which signal stimulus was "most similar" to Stimulus 2, the observer then remembered and gave that response consistently thereafter. It can be seen in Fig. 8 that the model accurately predicted the different distributions for the three observers -- but at a cost. The parameter-estimation algorithm has attempted to accommodate the extreme distribution by increasing the sensitivity parameter value in Condition 3 relative to Conditions 1 or 2 (see Table 5D), thus decreasing the error in predicting that particular response distribution. As a consequence, however, the

prediction of other response distributions are now more extreme than they otherwise would be. This effect can be seen in the predicted distributions for most other stimuli in Condition 3.

Over-all, however, the model, using the perceptual space derived from similarity judgments within the context of the full stimulus set, predicts quite well the confusion matrices for partial identification tasks in which different subsets of the stimuli are identified. The predictions are improved, as in the complete identification task, when a fourth dimension is added to the MDS-derived perceptual space.

Adaptive Tuning

Adaptive Tuning In Identification

Our application of the model to the complete identification task and the three conditions of the partial identification task suggests that the observers were flexible in their use of perceptual dimensions to identify the stimuli, in that our estimates of the relative salience weights on the several dimensions varied from condition to condition (see Fig. 9). Given that the same set of eight stimuli was presented in both tasks, we may ask what

motivated the observers to adjust the pattern of salience weights as they did across the different conditions. While the set of stimuli remained constant across all conditions, the subset of stimuli that we required the observer to identify -- the "signals" -- changed from condition to condition. Furthermore, it was only for this subset of stimuli that the observer received discriminative feedback that indicated which stimulus had been presented. We believe that the observer was engaged in an adaptive tuning process in which the relative weighting of dimensions was adjusted in order to maximize the discriminability of the subset of stimulus patterns to be identified in that condition. This tuning process probably takes place gradually, over many trials, based on the feedback given the observer regarding the correctness of identification.

Given that observers were instructed to maximize their percentage of correct identifications, it seems likely that this criterion formed their basis for tuning. To test this hypothesis, we have determined from the model, for each condition of the experiment, what pattern of dimension salience weights would maximize the probability of a correct signal identification. The optimal pattern of weights, assuming an average value for the sensitivity parameter, is plotted by the dashed line for each experimental condition

of both tasks in Fig. 9. While the detailed agreement of the observed and optimal weight patterns is not particularly good, there is a general correspondence of observed and predicted pattern shape across conditions, with the exception of Condition 3 of the 8 x 4 task. The comparison in this particular condition is probably not meaningful because of the large inter-observer differences, as discussed earlier. In general, the observed patterns of dimension salience weights seem consistent with the hypothesis that observers are tuning their weighting of dimensions in order to maximize the probability of a correct identification.

Adaptive Tuning In Judgment of Similarity

The concept of adaptive tuning may also provide an explanation for the failure of periodicity to emerge as a dimension in the INDSCAL analysis of the similarity judgments obtained from our three observers. If we assume that the observers' perceptual spaces were the same in both the similarity-judgment and identification tasks, then periodicity could fail to emerge as a dimension in the INDSCAL analysis if the observers were according it zero, or nearly zero, weight. This would be analogous to our analysis of Condition 2 of the partial identification task

in which dimension 3 (contrast) was given zero weight by all observers, although this dimension presumably was available since it was used in other conditions. Thus, we suggest that periodicity was present in the observers' perceptual spaces in the similarity-judgment task, but given zero weight in the adapted pattern of dimension weights. Moreover, we believe that the particular pattern of weights resulted from an adaptive tuning process that sought to optimize some aspect of performance, as in the identification task. Since observers were instructed to map their perceived range of similarities into numbers so as to use the entire range from one to ten, we speculate that observers may have tuned dimension weights to obtain the maximal possible range of inter-stimulus distances over all pairs of stimuli. Further work is required to decide this issue.

Discussion

A Validation of MDS Procedures and the Decision Model

Our approach in this paper assumes (1) that a set of complex stimuli can be represented perceptually as a set of points in a multidimensional psychological space, (2) that an MDS procedure can be used to derive the dimensions of that space and the relative loci of the stimuli within the

space, and (3) that identification confusions can be predicted by a simple decision model based on weighted interstimulus distances. Our success in predicting confusion matrices for individual observers across several experimental conditions provides support for this approach.

We believe that this outcome provides a significant validation of the use of MDS procedures to reveal perceptual structure. Other evidence for the validity of MDS-derived spaces has relied largely on the intuitive reasonableness of the abstracted dimensions and stimulus configurations. The present outcome provides much stronger support in that the MDS-derived space is used to predict data in a different, independent task. To be sure, the best prediction of identification behavior occurred when the MDS-derived space for our three observers was supplemented by an additional dimension. Nonetheless, the first three dimensions revealed by the MDS procedure for those observers were useful in the prediction.

Our results also support the decision model as a description of the identification process. The model accounts well for the changes in performance observed across different conditions of the complete and partial identification tasks. Of particular interest is the fact that the model accounts for these changes in performance in

terms of changes in the relative salience of perceptual dimensions rather than in terms of changes in the structure of the perceptual space. In the model, a given stimulus is assumed to have a fixed location in the perceptual space, regardless of changing stimulus context.

The MDS/Decision-Model Approach as an Aid in
Predicting Identification Performance

A new, particular identification task arises whenever new signals come under study. New signals are studied when new sources, sensors, settings, or displays are discovered, defined, or devised. Examples are replete in medical diagnosis, nondestructive testing of materials and equipment, military surveillance, analysis of biological microstructures, or study of the perceptually handicapped. Our experiments enhance the possibility that the present approach can be used to gain understanding about, and predict behavior in, any particular identification task.

Gaining understanding about a new set of signals, in this context, means isolating the perceptual dimensions that are useful, and assigning the proper saliences to those dimensions. The practical import of the approach described here is that perceptual dimensions can be isolated quite simply and quickly. We suspect that observers can be economically trained to use the useful dimensions, and to

use them approximately with weights that maximize the probability of correct identifications.

We have pointed out that our application of an MDS procedure to our three test observers did not define all of the dimensions that were found useful in the identification task, but that the additional dimension in question (periodicity) did appear in the four-dimensional MDS solution yielded by 14 other judges, some of whom had more general experience with signals of the sort used than the three observers. In examining a new identification task, one might do well to employ a rather large number of judges, having expertise with the general class of signals in question, and examine carefully the MDS solutions of larger dimensionality.

An alternative might be to ask the judges in the MDS task to rate the "confusability" of the stimulus pairs, with the identification task in their minds, rather than stimulus similarity. Our three test observers might have yielded the additional useful dimension (periodicity) under such instructions. Though our present experiments were constrained by our interest in validating MDS procedures as they are usually applied, MDS tasks undertaken specifically to analyze new sets of signals might possibly be adjusted in this manner to make them maximally useful.

References

- Carroll, J. D. Individual differences and multidimensional scaling. In R. N. Shepard, A. K. Romney, and S. Nerlove (Eds.). Multidimensional Scaling: Theory and Applications in the Behavioral Sciences. New York: Seminar Press, 1972, 105-155.
- Carroll, J. D., and Chang, J. J. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. Psychometrika, 1970, 35, 288-319.
- Carroll, J. D., and Wish, M. Models and methods for three-way multidimensional scaling. In Contemporary Developments in Mathematical Psychology. (Eds.) R. C. Atkinson, D. H. Krantz, R. D. Luce, and P. Suppes. San Francisco: Freeman, 1973, 57-105.
- Hardzinski, M., and Pachella, R. G. A psychophysical analysis of complex integrated displays. (Technical Report 59 (014523-2-T)). Ann Arbor, Michigan: University of Michigan, February, 1977.
- Howard, J. H. Jr. The psychophysical structure of eight complex underwater sounds. Journal of the Acoustical Society of America, 1977, 62, 149-156.
- Howard, J. H., Jr., and Silverman, E. G. A multidimensional scaling analysis of 16 complex sounds. Perception & Psychophysics, 1976, 19, 193-200.
- Klein, W., Plomp, R., and Pols, L. C. W. Vowel spectra, vowel spaces, and vowel identification. Journal of the Acoustical Society of America, 1970, 48, 999-1009.
- Luce, R. D. Detection and recognition. In Handbook of Mathematical Psychology. (Eds.) R. D. Luce, R. R. Bush, and E. Galanter. New York: Wiley, 1963, 103-189.
- Miller, J. R., and Carterette, E. C. Perceptual space for musical structures. Journal of the Acoustical Society of America, 1975, 58, 711-720.
- Morgan, B. J. T., Woodhead, M. M., and Webster, J. C. On the recovery of physical dimensions of stimuli, using multidimensional scaling. Journal of the Acoustical Society of America, 1967, 42, 773-777.

Pachella, R. G., and Somers, P. The development of integrated multidimensional displays. Naval Research Reviews (in press).

Plomp, R., and Steeneken, H. J. .M. Effect of phase on the timbre of complex tones. Journal of the Acoustical Society of America, 1969, 46, 409-421.

Pols, L. C. W., Van Der Kamp, L. J. Th., and Plomp, R. Perceptual and physical space of vowel sounds. Journal of the Acoustical Society of America, 1969, 46, 458-467.

Romney, A. K., Shepard, R. .N., and Nerlove, S. B. (Eds.). Multidimensional Scaling: Theory and Applications in the Behavioral Sciences, (Vol. 11). New York: Seminar Press, 1972.

Shaw, D. J. A phonological interpretation of two acoustic confusion matrices. Perception & Psychophysics, 1975, 17, 537-542.

Shepard, R. N. Psychological representation of speech sounds. In E. David and P. Denes (Eds.). Human Communication: A Unified View. New York: McGraw-Hill, 1972.

Shepard, R. N. Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. Psychometrika, 1957, 22, 325-345.

Shepard, R. N. Stimulus and response generalization: Deduction of the generalization gradient from a trace model. Psychological Review, 1958a, 65, 242-256.

Shepard, R. N. Stimulus and response generalization: Tests of a model relating generalization to distance in psychological space. Journal of Experimental Psychology, 1958b, 55, 509-523.

Shepard, R. N., and Chipman, S. Second-order isomorphism of internal representations: Shapes of states. Cognitive Psychology, 1970, 1, 1-17.

Shepard, R. N., Romney, A. K., and Nerlove, S. B. (Eds.) Multidimensional Scaling: Theory and Applications in the Behavioral Sciences, (Vol. 1). New York: Seminar Press, 1972.

Stenson, H. H. The psychophysical dimensions of similarity among random shapes. Perception & Psychophysics, 1968, 3, 201-114.

Swets, J. A., Green, D. A., Getty, D. J., and Swets, J. B. Signal detection and identification at successive stages of observation. (Technical Report 3535). Cambridge, Massachusetts: Bolt Beranek and Newman Inc., July, 1977.

Tversky, A., and Krantz, D. H. The dimensional representation and the metric structure of similarity data. Journal of Mathematical Psychology, 1970, 7, 572-596.

OFFICE OF NAVAL RESEARCH, CODE 455
TECHNICAL REPORTS DISTRIBUTION LIST

Director, Engineering Psychology Programs, Code 455 Office of Naval Research 800 North Quincy Street Arlington, Va 22217 (5cys)	Commanding Officer ONR Branch Office Attention: Dr. J. Lester 495 Summer Street Boston, Ma 02210
Defense Documentation Center Cameron Station Alexandria, Va 22214 (12 cys)	Commanding Officer ONR Branch Office Attention: Dr. Charles Davis 536 South Clark Street Chicago, Illinois 60605
Dr. Robert Young, Director Cybernetics Technology Office Advanced Research Projects Agency 1400 Wilson Blvd. Arlington, Va 22209	Commanding Officer ONR Branch Office Attention: Dr. E. Gloye 1030 East Green Street Pasadena, California 91106
Lt. Col. Henry I. Taylor USAF OAD (E&LS) ODDR&E Pentagon, Room 3D129 Washington, D. C. 20301	Dr. M. Bertin Office of Naval Research Scientific Liaison Group American Embassy, Room A-407 APO San Francisco 96503
Office of Naval Research International Programs Code 102IP 800 North Quincy Street Arlington, Va 22217 (6 cys)	Director, Naval Research Laboratory Technical Information Division Code 2627 Washington, D. C. 20375 (6 cys)
Director, Avionics & Weapons Programs, Code 212 Office of Naval Research 800 North Quincy Street Arlington, Va 22217	Mr. John Hill Naval Research Laboratory Code 5707.40 Washington, D. C. 20375
Director, Sensor Technology Program, Code 222 Office of Naval Research 800 North Quincy Street Arlington, Va 22217	Office of the Chief of Naval Operations, OP987P10 Personnel Logistics Plans Department of the Navy Washington, D. C. 20350

Report No. 3719

Bolt Beranek and Newman Inc.

Director, Physiology Program
Code 441
Office of Naval Research
800 North Quincy Street
Arlington, Va 22217

Commander
Naval Air Systems Command
Human Factors Programs, AIR340F
Washington, D. C. 20361

Commander
Naval Air Systems Command
Crew Station Design
AIR 5313
Washington, D. C. 20361

Mr. T. Momiyama
Naval Air Systems Command
Advance Concepts Division
AIR 03P34
Washington, D. C. 200361

Commander
Naval Electronics Systems Command
Human Factors Engineering Branch
Code 4701
Washington, D. C. 20360

ICDR T. W. Schropp
Naval Sea Systems Command
NAVSEA 00C3
Washington, D. C. 20362

Mr. James Jenkins
Naval Sea Systems Command
NAVSEA
Code 06H1-3
Washington, D. C. 20362

Mr. Arnold Rubinstein
Naval Material Command
NAVMAT 0344
Department of the Navy
Washington, D. C. 20360

Director
Behavioral Sciences Department
Naval Medical Research Institute
Bethesda, MD 20014

Dr. George Moeller
Human Factors Engineering Branch
Submarine Medical Research Lab.
Naval Submarine Base
Groton, Connecticut 06340

Chief,
Aerospace Psychology Division
Naval Aerospace Medical Institute
Pensacola, Fla 32512

Navy Personnel Research and
Development Center
Management Support Department
Code 210
San Diego, California 92152

Dr. Fred Muckler
Navy Personnel Research and
Development Center
Manned Systems Design, Code 311
San Diego, California 92152

ICDR Michael O'Bar
Navy Personnel Research and
Development Center
Code 305
San Diego, California 92152

Report No. 3719

Bolt Beranek and Newman Inc.

Dr. James Curtin
Naval Sea Systems Command
Personnel & Training
Analyses Office
NAVSEA 074C1
Washington, D. C. 20362

LCDR R. Gibson
Bureau of Medicine & Surgery
Aerospace Psychology Branch
Code 513
Washington, D. C. 20372

CDR Paul Nelson
Naval Medical R&D Command
Code 44, Naval medical Center
Bethesda, MD 20014

Mr. Ronald A. Erickson
Human Factors Branch
Code 3175
Naval Weapons Center
China Lake, California 93555

Human Factors Section
Systems Engineering Test
Directorate
Patuxent River, MD 20670

Human Factors Division
Naval Ocean Systems Center
San Diego, California 92152

Human Factors Engineering Branch
Naval Ship Research and
Development Center
Annapolis Division
Annapolis, MD 21402

Dr. Robert French
Naval Ocean Systems Center
San Diego, California 92132

Dr. Jerry C. Lamb
Display Branch, Code TD112
Naval Underwater Systems Center
New London, Connecticut 06320

Naval Training Equipment Center
Attention: Technical Library
Orlando, Florida 32812

Mr. A. V. Anderson
Naval Personnel Research and
Development Center
Code 302
San Diego, California 92152

Human Factors Engineering Branch
Crew Systems Department
Naval Air Development Center
Johnsville
Warminster, Pa 18974

Dr. Gary Poock
Operations Research Development
Naval Postgraduate School
Monterey, California 93940

Dr. A. I. Slafkosky
Scientific Advisor
Commandant of the Marine Corps
Code RD-1
Washington, D. C. 20380

Mr. J. Barber
Headquarters, Department of the
Army, DAPE-PBR
Washington, D. C. 200546

Dr. Joseph Zeidner, Director
Organization and Systems Research
Laboratory
U. S. Army Research Institute
1100 Wilson Boulevard
Arlington, Va 22209

Technical Director
U. S. Army Human Engineering Labs.
Aberdeen Proving Ground
Aberdeen, MD 21005

U. S. Air Force Office of
Scientific Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, D. C. 20332

Dr. Donald A. Topmill
Chief, Systems Engineering Branch
Human Engineering Division
USAF AMRL/HES
Wright-Patterson AFB, Ohio 45433

Report No. 3719

Bolt Beranek and Newman Inc.

Human Factors Department
Code N215
Naval Training Equipment Center
Orlando, Florida 32813

Dr. Alfred F. Smode
Training Analysis & Evaluation
Group, Naval Training
Center, Code N-00T
Orlando, Fla 32813

Dr. W. S. Vaughan
Oceanautics, Inc.
3308 Dodge Park Road
Landover, MD 20785

Dr. Arthur I. Siegel
Applied Psychol. Services, Inc.
404 East Lancaster Street
Wayne, Pa 29087

Dr. Robert R. Mackie
Human Factors Research Inc.
Santa Barbara Research Park
6780 Cortona Drive
Goleta, California 93017

Dr. Gershon Weltman
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, Calif. 91364

Dr. J. W. Wulfeck
New Mexico State University
Department of Psychology
Box 5095
Las Cruces, New Mexico 88003

Air University Library
Maxwell Air Force Base
Alabama 36112

Dr. Robert Williges
Human Factors Laboratory
Virginia Polytechnic Institute
130 Whittemore Hall
Blacksburg, Va 24061

Dr. James H. Howard
Catholic University
Department of Psychology
Washington, D. C. 20064

Dr. G. H. Robinson
University of Wisconsin
Department of Industrial Eng.
1513 University Avenue
Madison, Wisconsin 53706

Mr. E. M. Connelly
Omnemii, Inc.
410 Pine Street, S. E.
Suite 200
Vienna, Va 22180

Dr. Thomas L. Harrington
University of Nevada
Department of Psychology
Reno, Nevada 89507

Dr. Robert G. Pachella
University of Michigan
Department of Psychology
330 Packard Road
Ann Arbor, Michigan 48104

Report No. 3719

Bolt Beranek and Newman Inc.

Dr. H. Rudy Ramsey
Science Applications, Inc.
40 Denver Tech Center West
7935 East Prentice Avenue
Englewood, Colorado 80110

Dr. Paul Slovic
Decision Research
1201 Oak Street
Eugene, Oregon 97401

Dr. Ross I. Pepper
Naval Undersea Center
Hawaii Laboratory
P. O. Box 997
Kailua, Hawaii 96734

Journal Supplement Abstract
Service
American Psychological Assoc.
1200 17th Street, N. W.
Washington, D. C. 20036 (3 cys)

Director, Human Factors Wing
Defense & Civil Institute
of Environmental
Medicine, P.O. Box 2000
Downsville, Toronto
Ontario, CANADA

Dr. David Zaidel
Road Safety Centre
Technion City
Haifa, ISRAEL

Prof. Dr. Carl Graf Hoyos
Institute for Psychology
Technical University
8000 Munich
Arcisstr 21
FEDERAL REPUBLIC OF GERMANY

Dr. Robert Fox
Vanderbilt University
Department of Psychology
Nashville, Tennessee 37240

Dr. Jo Ann S. Kinney
Submarine Medical Research Lab.
Naval Submarine Base
Groton, Connecticut 06340

Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army-Navy Drive
Arlington, Va 222002

Dr. Stanley Deutsch
Office of Life Sciences
Headquarters, NASA
600 Independence Avenue
Washington, D. C. 20546

R. W. Bryant
ASW-132, ASWPO
NAVSEA 2W28
National Center No. 1
Washington, D. C. 20360

Dr. A. D. Baddeley
Applied Psychology Unit
Medical Research Council
15 Chaucer Road
Cambridge, CB2 2EF
ENGLAND